PHILOSOPHY FOR THE GLOBAL WORLD

PREFACE

This text grew out of a set of PowerPoint presentations—nearly 450 detailed and increasingly elaborate slides—that I developed over many years for the *Global Justice* course I have taught since 2016 at the University of Rome Tor Vergata. A shorter version of the course was also offered for few years at the Summer School of the China University of Political Science and Law (CUPL) in Beijing. Over time, the course evolved through dialogue: through students' questions, objections, confusions, and insights. I owe a special debt of gratitude to the students of Global Governance at Tor Vergata and to those at CUPL, whose enthusiasm and curiosity continually pushed me to refine my explanations and broaden the conceptual scope of the lectures. Over the years I also supervised numerous bachelor's theses on specific themes from the course, and these projects—thanks to the dedication of the students—expanded my own understanding and contributed to the ongoing reshaping of the course itself.

At a certain point, the gap between my ever-evolving lectures and the texts originally assigned in the syllabus became too wide to ignore. I had long used Michael Sandel's *Justice: What's the Right Thing to Do?* alongside Kok-Chor Tan's *What Is This Thing Called Global Justice?* as the closest approximation to the type of course I envisioned. Sandel's book introduces foundational questions in moral and political philosophy, though it does not engage global justice directly. Tan's book, by contrast, provides an accessible introduction to the central themes of global justice—poverty, migration, borders, climate change—but without fully covering the broader theoretical architecture I aimed to build. Together, these works served as a starting point, but they no longer captured the structure or ambition of the course.

This book is therefore an attempt to gather, integrate, and consolidate those years of teaching into a single, coherent text. It is written *for students*—and thus adopts a pedagogical structure—yet it preserves an academic tone and analytical rigor. The audience of the course is deeply international: in a typical cohort of more than one hundred students, more than twenty countries are represented. Most have little or no prior training in philosophy, except for Italian students coming from classical or scientific high schools. However, by the time they take *Global Justice*, all students have been exposed to *Logic and Philosophy of Science* and to *Microeconomics*, both of which provide useful foundations. Still, no formal prerequisites are required. Technical points are placed in dedicated **Boxes**, and more demanding philosophical material is offered in **Appendices** at the end of each chapter, so that students can approach the core content without being overwhelmed.

The structure of the book reflects this dual ambition: conceptual clarity and theoretical depth, paired with empirical relevance.

- **Part I** lays the conceptual foundations. Before engaging with the major theories of social ethics and justice —utilitarianism, libertarianism, contractualism, communitarianism—it is essential to clarify the basic moral grammar that underlies them: duties and rights, self-interest and altruism, individual freedom and collective responsibility. These ideas form the essential vocabulary for everything that follows.
- **Part II** presents the central theories of social ethics and justice. Here the book follows, in part, the classification suggested by Michael Sandel in *Justice: What's the Right Thing to Do?* while applying it to the global context. This is the most philosophical and architectonic part of the work: it surveys the great normative frameworks that have shaped modern debates about justice.
- **Part III** turns from theory to application. It examines how ethical principles intersect with economics and public policy in addressing poverty, migration, and climate change—three of the defining challenges of the twenty-first century. This is where abstract moral arguments must meet institutional design, empirical data, and political constraints.

The aim of *Philosophy for the Global World* is not merely to introduce ideas, but to show why they matter. The global world confronts us with moral questions that cannot be answered by national traditions alone. Climate change, financial interdependence, global supply chains, pandemics, human mobility, and technological disruption demand new ways of thinking about justice—ways that recognize the depth of our interdependence and the fragility of our shared planet.

This book is offered in that spirit: as an invitation to think philosophically about the world we inhabit, and about the institutions we might yet create to make it more just.

Sommario

PART I

Chapter 1

DUTY, RIGHTS, AND IDENTITY

1.1 Introduction

This chapter examines the moral foundations of justice in a world characterized by deep interdependence, persistent inequality, and expanding forms of collective power. It starts from a simple but often overlooked insight: moral and political claims cannot be adequately understood if rights are treated as self-standing entitlements detached from the obligations that make them effective. Rights acquire practical and normative force only insofar as individuals and institutions recognize and discharge the duties that correspond to them. To understand justice, therefore, we must begin not with what people are entitled to, but with what is owed.

The chapter develops this duty-centered perspective by exploring how moral obligations are structured, justified, and transformed as social life becomes more complex. While traditional moral thought often distinguishes sharply between negative duties of non-interference and positive duties of assistance, contemporary conditions blur this boundary. Many of the most serious harms individuals face today are not the result of isolated actions but of participation in large-scale economic, political, and environmental systems. In such contexts, avoiding harm increasingly requires active forms of restraint, regulation, and institutional design. Moral responsibility thus extends beyond personal conduct to collective practices and shared structures.

At the same time, moral agency is never exercised in abstraction. Human beings are embedded in relationships, communities, and identities that shape their motivations, allegiances, and sense of responsibility. The chapter therefore situates universal moral obligations within the realities of partial attachment, recognition, and belonging. Rather than treating these features of moral life as obstacles to universality, it argues that they are among the conditions through which universal principles become practically meaningful. Moral impartiality must be compatible with the fact that obligations are lived and enacted through particular social and institutional contexts.

Building on this framework, the chapter turns to the problem of justice at the global level. As political and economic institutions increasingly operate across borders, the traditional assumption that duties of justice are confined to sovereign states becomes difficult to sustain. Global trade regimes, financial systems, migration controls, and environmental governance structures exert coercive influence over individuals who lack meaningful control over their design. Where such institutions foreseeably and avoidably undermine access to basic goods, moral responsibility cannot be reduced to voluntary charity. It takes the form of duties of justice grounded in the same logic that underpins domestic rights.

The central claim of the chapter is that justice must be understood as an institutional achievement rooted in moral obligation. Rights to security, subsistence, and freedom are not merely declared; they are secured through sustained collective action. In an interconnected world, this action must extend beyond national borders, even in the absence of a global sovereign. By foregrounding the priority of duties over rights, the chapter offers a framework for assessing both domestic and global institutions according to their impact on the basic conditions of human agency and dignity. Justice, on this view, is not a matter of ideal aspiration but of responsibility for the social structures through which human lives are shaped.

1.2 Morality, Self-Interest, and Collective Action: Gandhi, Smith, Sidgwick, and the Limits of Ethical Egoism

Morality is commonly understood as requiring impartiality, concern for others, and adherence to principles that extend beyond personal advantage. By contrast, self-interest directs agents toward the promotion of their own good. While these two orientations often appear to conflict, many philosophers have argued that self-interested behavior can, under certain conditions, produce socially beneficial outcomes. This raises a deeper philosophical problem: can self-interest itself serve as a foundation for moral action, or does genuine morality necessarily require transcending it?

Mahatma Gandhi grounds morality in the intention behind action rather than in its external effects. He insists that a moral act must arise from free and conscious willing:

“A moral act must be our own act; it must spring from our own will. If we act mechanically, there is no moral content in our act.” (Gandhi (1968), p.5)

For Gandhi, the moral worth of an action depends not merely on its outward effect but on the motivation that underlies it. He continues:

“It is not enough that an act done by us is in itself good; it should have been done with the intention to do good. Two men may have done exactly the same thing; but the act of one may be moral, and that of the other the contrary.” (Gandhi 1968, p.5)

The example Gandhi provides — one man feeding the poor out of compassion and another doing so to gain prestige — illustrates that morality requires freedom from coercion and self-interest. Acting morally is not simply a matter of performing socially approved acts; it involves **the conscious choice to do good for its own sake**, free from ulterior motives. In this framework, acting selfishly is inherently at odds with true moral action, because self-interest contaminates the purity of intention. Genuine moral action requires freedom from coercion and ulterior motives, and thus cannot be grounded in egoistic reasoning. On this view, acting selfishly and acting morally are fundamentally incompatible.

Where Gandhi emphasizes moral intention, Adam Smith and Bernard Mandeville focus on the social consequences of self-interested behavior. Smith, in *An Inquiry into the Nature and Causes of the Wealth of Nations (1776/1976),* famously observes:

"It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own self-interest. We address ourselves not to their humanity but to their self-love. (pp. 26-27)"

And Bernard Mandeville goes further, in the Preface to the 1714 edition of *The Fable of the Bees* writes (see BOX # 1.1):

"The Moment Evil ceases, the Society must be spoiled, if not totally destroyed".

If individuals stopped acting form self-interest –seeking luxury, competition, or pride – economic and cultural vitality would collapse. What looks morally corrupt on the individual level may sustain prosperity at the social level.

These arguments show that self-interest can function as an effective organizing principle for economic life. Through market mechanisms, individual pursuits of gain may unintentionally generate collective benefits. However, this does not amount to a moral justification of self-interest. Smith does not claim that self-interested actions are morally virtuous, only that they can be socially useful. The invisible hand explains coordination, not moral worth. Thus, while self-interest may produce good outcomes, it does not satisfy Gandhi's requirement that moral action be motivated by concern for others rather than by personal advantage.

Henry Sidgwick confronts directly the question of whether self-interest can serve as a moral principle. In *The Methods of Ethics*, he distinguishes between **egoistic hedonism**, which holds that each person ought to maximize their own good; **utilitarianism**, which seeks to maximize overall good; and i**ntuitionism**, which relies on self-evident moral duties. Sidgwick takes ethical egoism seriously as a normative view, recognizing its intuitive appeal and its alignment with prudence and rational self-concern.

However, Sidgwick ultimately rejects egoism as a moral theory. While it may be rational from the first-person perspective to prioritize one's own good, egoism fails to justify this priority from an **impartial** standpoint. There is no rational reason why one individual's happiness should count for more than another's. This leads to what Sidgwick famously describes as a "dualism of practical reason": a conflict between the dictates of prudence and the demands of morality.

Ethical egoism fails to meet the basic conditions of a moral theory. First, it lacks **universality**: it issues different and potentially incompatible prescriptions to different agents. Second, it provides no rational method for resolving conflicts of interest, whereas morality aims precisely to regulate such conflicts through shared principles.

Third, it undermines collective rationality, since actions that are individually rational often produce outcomes that are worse for everyone involved.

The structural weaknesses identified by Sidgwick become especially clear in game-theoretic contexts. In the Prisoner's Dilemma, each agent has a rational incentive to defect, yet mutual defection produces a worse outcome than mutual cooperation (See BOX # 1.2 for details). .

Similarly, in public goods dilemmas analyzed by Derek Parfit in *Reasons and Persons* (1984), individually rational behavior leads to collectively disastrous results:

- **Polluters:** Each benefits for non bearing the costs of non-polluting, if all do so, pollution will prevail.
- **Commuters**: Each benefits personally by driving faster, yet if all do so, traffic congestion worsens.
- **Fishermen**: Overfishing maximizes short-term gain for individuals, but collective overfishing depletes the resource.
- **Soldiers in the battle**: Acting purely in self-interest to save her/his own life may seem rational, but if all follow the same logic, the battle is lost.

In public goods dilemmas, individual rationality often conflicts with collective welfare. Self-interested behavior can undermine outcomes that would be better for everyone. This exposes a paradox: **Egoism is self-defeating in several circumstances at the collective level.**
Rational self-interest can lead to collectively irrational outcomes. Thus, even from a prudential standpoint, egoism undermines the very conditions of individual well-being. Egoism reduces morality to calculation: what benefits *me*?

But moral action often requires *sacrifice* for others or adherence to principles even when self-interest suffers. For instance, telling the truth, keeping promises, or helping a stranger in distress may not always serve one's immediate welfare. Yet we still regard such acts as morally praiseworthy. Egoism cannot explain why we admire self-sacrifice or impartial justice—except by redefining them in self-serving terms, which empties them of genuine moral meaning. This is why philosophers like Kant and Singer, although from different perspectives, argue that **true morality begins when we transcend self-interest**. The egoistic standpoint is **morally non-reciprocal**: it asks others to act morally only when it benefits oneself, which violates the principle of universalizability.

The perspectives of Gandhi, Smith, Sidgwick, and Parfit collectively reveal the conceptual and practical limitations of grounding morality in self-interest. Gandhi emphasizes the indispensability of moral intention and the incompatibility of genuine moral action with egoistic motives. Smith and Mandeville show that self-interest can generate socially beneficial outcomes, but these benefits do not impart moral worth. Sidgwick exposes the structural incoherence of ethical egoism by demonstrating its inability to satisfy the demands of universality and its propensity to produce collective irrationality. Parfit's analysis of public goods dilemmas reinforces and extends this critique, illustrating the real-world consequences of unrestrained self-interest.

Taken together, these accounts underscore that while self-interest may align with social welfare under particular conditions, it cannot function as the foundation of morality. Moral action requires a standpoint that transcends personal advantage—one grounded in impartiality, reciprocity, and a recognition of the interdependence that characterizes human life

Box # 1.1: The Fable of the Bees

Bernard Mandeville's *The Fable of the Bees* (1714) is a satirical poem and philosophical commentary that argues that private vices—such as greed, selfishness, ambition and luxury—can lead to public benefits, such as economic prosperity and social advancement. Mandeville challenges the traditional moral belief that virtue is essential for a thriving society. He suggests that selfish behaviors, like the pursuit of wealth and luxury, drive economic activity, industry, and innovation. He exposes how societies condemn vices while benefiting from them. For example, luxury spending may seem wasteful, but it creates jobs and sustains economies. His ideas foreshadow Adam Smith's *invisible hand* theory, as they suggest that individuals pursuing their self-interest inadvertently contribute to the common good. In the poem, when the bees become virtuous and abandon their greed and ambition, their once-thriving hive collapses. This illustrates Mandeville's belief that a society purely based on virtue lacks the energy and motivation to progres. Mandeville's ideas were controversial because they challenged moral and religious doctrines of his time. However, his work influenced economic thought, particularly classical liberalism and capitalism ethics.

Box # 1.2: The Prisoner's Dilemma

Two political prisoners are brought before a magistrate, who presents them with the following "offer." If one prisoner confesses and testifies against the other, the informant will go free while the other receives a ten-year sentence. If both betray each other, each will be sentenced to five years. However, if both remain silent, each will serve only one year. The prisoners are then placed in separate rooms to deliberate their decisions. They are assured that their choices will remain secret, that they will never meet again, and that their decisions will have no further consequences beyond the sentences described. This is the so-called *Prisoner's Dilemma*, a thought experiment much discussed by economists, political scientists, and moral philosophers. Several important insights arise from this thought-experiment (cells identified like directions on a map):

	Prisoner B: Silent	**Prisoner B: Confess**
Prisoner A: Silent	Both serve 1 year (NW)	A serves 10 years, B goes free (NE)
Prisoner A: Confess	A goes free, B serves 10 years (SW)	Both serve 5 years (SE)

Several lessons emerge from this model:

1. **Interdependence of outcomes:** The fate of each prisoner is not solely in his own hands. Each is dependent on the other's choice, and vice versa. An individualistic strategy is self-defeating: if Prisoner B chooses "Confess" (betray), Prisoner A's best option is also to "confess" (betray). If B chooses "silent" (cooperate) A still maximizes his individual gain by choosing "confess" (betray). Thus, regardless of B's decision, A will always betray — and the same reasoning applies to B. This is the so-called *Nash equilibrium*.
2. **Conflict of rationalities:** The optimal individual outcome (cells SW and NE: one prisoner goes free, the other gets ten years) is not the optimal collective outcome (cell SE: two years total, one year each).
3. **The paradox of self-interest:** Pursuing one's own maximum benefit leads both to worse results. Egoism, when universalized, becomes self-destructive.
4. **The moral or cooperative outcome:** The optimal collective solution (one year each) depends on trust, altruism, and reciprocity — moral dispositions absent from pure egoism.
5. **Social enforcement and moral order:** When rules, sanctions, or reputational incentives are introduced, cooperation becomes the rational choice. Law and morality thus function to align self-interest with the common good — a solution Sidgwick sought through rational utilitarianism.

Sidgwick recognized this contradiction at the heart of egoism: while self-interest can justify moral restraint when aligned with long-term benefit, it cannot rationally explain why an individual should sacrifice for others when no advantage can be gained. This is what he called the "dualism of practical reason." The *Prisoner's Dilemma* dramatizes precisely this dualism — the rational egoist's pursuit of self-interest undermines the very conditions that make cooperation and trust possible.

In this sense, the dilemma exposes the limits of egoism as a moral method. Rational egoism cannot generate the moral bonds it depends upon; it must rely on external structures — social norms, legal sanctions, or shared ethical principles — to transform individual prudence into collective morality. Only when self-interest is redefined through mutual recognition, as in Sidgwick's synthesis of reason and benevolence, does moral cooperation become both rational and just.

1.3 Adam Smith on Moral Sentiments and the Foundations of Justice

Adam Smith's first major work, *The Theory of Moral Sentiments* (1759/1976) (TMS), is not a treatise of moralism prescribing rules for virtuous living; rather, it is a pioneering work of **moral psychology**, aimed at understanding how human beings develop moral sentiments and how these sentiments evolve over time. Smith's approach is empirical, in the fashion of Newton, seeking to observe, describe, and explain patterns in human behavior rather than dictate abstract moral imperatives. He is interested in how individuals interact, respond, and adjust to one another, and how such interactions give rise to shared moral standards.

A central claim in Smith's moral psychology is that humans **desire a mutual sympathy of sentiments**. We experience pleasure when our feelings are reflected and shared by others, and disquiet when they are not. This desire for sympathy has profound social consequences. Living in a community, individuals gradually develop **standards of behavior** through repeated interaction. This process of mutual adjustment gives rise to shared habits, customs, and moral rules — ranging from etiquette to moral duties. Smith conceptualizes this process through the notion of the **"impartial spectator"**, an imagined perspective that allows individuals to judge their own conduct and that of others, facilitating social harmony and moral self-regulation.

Smith emphasizes that **moral sentiments evolve through social interaction**. Humans are intrinsically social: a solitary individual, as in Smith's thought experiment of the isolated islander, would lack the necessary social feedback to develop moral sentiments. Moral awareness depends on being mirrored by others; it is through engagement in society that we acquire the capacity to desire not only to be loved but also to be **lovely**, that is, worthy of affection and respect. As Smith observes:

"The chief part of human happiness arises from the consciousness of being beloved. (...) All the members of human society stand in need of each other's assistance, and are likewise exposed to mutual injuries. When the necessary assistance is reciprocally afforded from love, from gratitude, from friendship, and esteem, the society flourishes and is happy.(TMS, II.i.1.1, p. 9.)."

This social dimension also accounts for both moral variation and consistency. While moral sentiments differ across cultures — reflecting **moral relativism** — certain fundamental sentiments overlap universally, suggesting a degree of **moral objectivity**.

This duality allows Smith to explain both the evolution of ethical norms over time and the persistence of foundational moral principles, such as justice and benevolence. Over the past decades, for example, societal attitudes toward same-sex marriage, divorce, or ethnic and sexual equality have shifted dramatically. Smith's framework anticipates such changes: moral sentiments are not fixed, but evolve as society evolves.

This empirically grounded understanding of moral psychology provides the foundation for Smith's later discussion of **negative and positive virtues**, which underpin his account of justice and the duties required for social life.

Smith's distinction between **negative** and **positive virtues** is central to his moral theory. In *TMS*, he defines **justice** as a virtue that "does real and positive hurt to some particular persons" when violated, and therefore "is capable of being restrained by the interposition of the magistrate (*TMS*, II.ii.1.5, p. 78)". Justice concerns actions whose omission or commission directly affects others' security or rights. By contrast, **beneficence** "is always free; it cannot be extorted by force, the mere want of it exposes to no punishment, because the mere want of beneficence tends to do no real positive evil (*TMS*, II.ii.1.6, p. 78)."

Justice is **negative** because it consists in *refraining* from causing harm, whereas beneficence is **positive** because it requires *active* engagement in promoting the good of others. Smith captures this difference with a famous analogy: "The rules of justice may be compared to the rules of grammar; the rules of the other virtues, to the rules which critics lay down for the attainment of what is sublime and elegant in composition. What is agreeable in the one is a matter of strict obligation; what is laudable in the other is the object only of praise (*TMS*, II.ii.1.9, p. 79)." Justice is necessary for social coherence; beneficence is admirable but non-compulsory.

Smith articulates justice through **three basic rules**, summarized as the "three Ps" — *person, property, and promises*. These rules express the moral boundaries that define just conduct:

1. **Guard the life and person of our neighbour** — Do not harm, kill, enslave, or molest.
2. **Guard our neighbour's property and possessions** — Do not steal, trespass, or damage another's goods.
3. **Guard our neighbour's rights, or what is due to him from the promises of others** — Do not renege on contracts or promises.

These rules correspond to **negative virtues**: obligations to refrain from harming others. As Smith observes, "We may often fulfil all the rules of justice by sitting still and doing nothing (*TMS*, II.ii.1.9, p. 79)." Justice thus requires restraint rather than action. Society "may subsist, though not in the most comfortable state, without beneficence; but the prevalence of injustice must utterly destroy it (*TMS*, II.ii.3.3, p. 86)." Justice forms the indispensable foundation of social life, while beneficence enhances moral excellence but is not strictly required.

Smith's negative and positive virtues closely parallel the modern distinction between **negative** and **positive duties**. *Negative duties* are obligations to refrain from harming others (e.g., killing, stealing, violating rights). *Positive duties*are obligations to promote the welfare of others (e.g., helping those in need, contributing to public goods). Kant formalized this as **perfect** (enforceable) versus **imperfect** (discretionary) duties, while contemporary rights theorists distinguish **duties of non-interference** from **duties of aid**.

Smith's negative virtue of justice aligns with **perfect, negative duties**: strict, enforceable obligations necessary for social order. Positive virtues, such as beneficence, correspond to **imperfect, positive duties**: morally praiseworthy but not enforceable.

In Smith's political thought, government exists primarily to enforce justice by protecting individuals' *person, property, and promises*. The state intervenes only when these are violated, reflecting a **"Negative Defensive Only"** conception of justice. Its role is to protect citizens from harm and maintain public order.

However, in *The Wealth of Nations*, Smith recognizes that government has additional responsibilities: "the duty of erecting and maintaining certain public works and certain public institutions which can never be for the interest of any individual, or small number of individuals, to erect and maintain (WN, V.i.c.1–2.)." Theese include **infrastructure**, **education**, and other **public goods** necessary for society. Even here, such duties are exceptions justified by utility rather than expressions of distributive justice in the modern sense.

Smith's account of justice — grounded in *negative duties* protecting person, property, and promises — provides a rigorous foundation for social order. Yet, by modern standards, it is **"too thin"**, since it excludes positive duties of assistance or redistribution. The state enforces only the limits of justice, acting when rights are infringed but not when needs are unmet. Nevertheless, Smith's acknowledgment of public works shows that even within his framework, justice extends slightly beyond strict non-injury.

Modern conceptions of **social justice** expand upon Smith's negative defensive model, including at least some **positive duties of beneficence**, reflecting obligations to assist those in need. Smith thus offers both a foundation and a contrast: the duty not to harm is essential, but contemporary moral thought recognizes that a full conception of justice also requires the duty to promote well-being,

1.4 Adam Smith on Self-Interest, Benevolence, and the Impartial Spectator: A Contrast with Mandeville

The writings of Bernard Mandeville and Adam Smith reveal two sharply different understandings of human motivation. Mandeville, in The Fable of the Bees, famously argued that private vices—selfish passions and vanity—could unintentionally produce public benefits. For him, self-interest was not only inevitable but also the hidden engine of social prosperity (see BOX n.2). Adam Smith, though often portrayed as a champion of self-interest, offers a much more complex picture. In TMS, he insists that both self-interest and benevolence are vital and complementary forces shaping human behavior.

Smith begins from the observation that human beings naturally share in the feelings of others: when we recognize that something makes another person unhappy, that recognition makes us, in turn, unhappy; the same holds for happiness. Yet Smith is clear that people are not motivated by benevolence alone. We are capable of being self-interested while at the same time holding other concerns. Self-interest and benevolence, for Smith, are not opposing impulses but a pair of interwoven and mutually reinforcing aspects of human nature—something he expects his readers to understand.

Nevertheless, Smith is a careful realist about the limits of sympathy. When it comes to matters of personal well-being—health, wealth, or safety—our own small interests typically weigh far more heavily than the greatest concerns of strangers. He illustrates this with a striking thought experiment in TMS:

> "Let us suppose that the great empire of China, with all its myriads of inhabitants, was suddenly swallowed up by an earthquake, and let us consider how a man of humanity in Europe, who had no sort of connexion with that part of the world, would be affected upon receiving intelligence of this dreadful calamity. He would, I imagine, first of all, express very strongly his sorrow for the misfortune of that unhappy people, he would make many melancholy reflections upon the precariousness of human life, and the vanity of all the labours of man, which could thus be annihilated in a moment. He would too, perhaps, if he was a man of speculation, enter into many reasonings concerning the effects which this disaster might produce upon the commerce of Europe, and the trade and business of the world in general. And when all this fine philosophy was over, when all these humane sentiments had been once fairly expressed, he would pursue his business or his pleasure, take his repose or his diversion, with the same ease and tranquillity, as if no such accident had happened. The most frivolous disaster which could befal himself would occasion a more real disturbance." (TMS, Part III, ch. 3, §4)

Smith presses the example further. If the same man were told that he would lose his little finger the next day, he would likely lie awake all night, troubled by the prospect. Yet the knowledge of a hundred million deaths in China would not keep him from sleeping soundly. For Smith, this stark contrast rests on two basic facts: first, the natural primacy of self-interest, and second, the tendency of human beings to feel more strongly attached to their own society than to distant strangers. Benevolence, he argues, is strongest within the family and diminishes as we move outward. Beyond the circle of kin and friends, we often find not merely indifference but, at times, active hostility. Sympathy toward strangers, therefore, is inevitably weaker.

But Smith does not end with this bleak assessment. He immediately asks whether a man of humanity, if given the choice to save those hundred million lives at the cost of his own little finger, would refuse the sacrifice. His answer is emphatic: no. While our **passive** feelings are indeed selfish, our **active** moral principles are often generous. A person could not bear to live with himself if he chose to save his finger and let millions perish. The decisive motive here is not sentimental love for the Chinese, for whom he may feel little or no affection, but rather a concern for the dignity and integrity of his own character. He must appear worthy in his own eyes.

This is the point at which Smith introduces his celebrated concept of the **impartial spectator**—the internalized figure of an ideal observer who represents the moral judgments of society. Each of us, Smith argues, learns to see ourselves through the eyes of others. Society acts like a mirror: our actions are constantly rehearsed before an imagined audience, and the desire to win the approval of this audience disciplines our conduct. We not only crave the approval of others; we also wish to be the kind of person who deserves approval.

The impartial spectator thus embeds self-interest within a moral framework of social reciprocity. To act well is to satisfy an inner standard that reflects the collective conscience. Smith's conception of self-interest, therefore, is not an invitation to unrestrained egoism. Rather, the liberty on which commercial society rests presupposes a moral order—one that links the individual's desire for esteem to the norms and codes of conduct that make social life possible.

In contrast to Mandeville's provocative claim that private vices yield public benefits without the guidance of moral sentiment, Smith offers a subtler vision. He acknowledges the undeniable force of self-interest but shows how it is checked and shaped by sympathy, the impartial spectator, and the desire to be worthy of admiration. Human beings, for Smith, are not merely self-regarding creatures; they are moral agents whose very self-interest is inseparable from their social and ethical nature.

1.5 Moral Preference, Allegiance, and the Expanding Circle

The question of how we ought to weigh our moral preferences and allegiances toward others lies at the heart of moral philosophy. Our moral intuitions tell us that while we recognize universal duties not to harm others, our positive duties — to aid, to care, to benefit — seem to vary according to our proximity or emotional attachment to those around us. We feel stronger obligations toward our family, friends, or compatriots than toward distant strangers as Adam Smith has pointed out. Yet, as Peter Singer argues in *The Expanding Circle* (1981), the development of moral reasoning pushes us beyond such local allegiances toward an ever-wider circle of concern, ultimately embracing all sentient beings. The challenge, then, is to reconcile our **partial attachments** — which are psychologically natural and socially embedded — with the **impartial demands of moral reason**.

Common sense morality, or what Henry Sidgwick called "the morality of common sense" in *The Methods of Ethics* (1874), acknowledges that moral life is structured by both **negative** and **positive duties**. Negative duties — such as not to lie, steal, or kill — are generally universal; they apply equally to friends and strangers, fellow citizens and foreigners. By contrast, positive duties — acts of benevolence or charity — are often guided by personal ties and emotional allegiances. It seems morally natural to save one’s child before saving a stranger’s, or to prioritize the welfare of one’s nation before that of others.

However, this intuitive partiality raises a normative question: *should* we trust our moral intuitions? As reflective beings, we not only ask *what* we ought to do, but *why* we ought to do it. The fact that certain partialities feel natural does not automatically make them morally justified. Moral philosophy, from Kant to Singer, seeks to examine the **reasons** behind these moral distinctions.

Singer begins from the evolutionary insight that morality has **biological origins**. Humans, like other social animals, evolved altruistic behaviors toward kin and members of their group because these behaviors enhanced genetic survival. Early altruism was thus partial and reciprocal — directed toward those who could return favors or shared common genes.

However, what distinguishes humans is **reason**. Rational reflection allows us to see that the suffering or happiness of others cannot be morally discounted merely because they belong to another family, tribe, or nation. As Singer writes:

“Reason is what allows us to see that the suffering of others matters as much as our own, and that there is no rational basis for restricting our moral concern to those near to us.”

This reflective process produces an **expanding circle** of moral concern: from family, to tribe, to nation, to all humanity, and eventually to all sentient beings. Rational impartiality, therefore, is not in conflict with human nature but represents the natural extension of our evolved capacities for empathy and reasoning.

The virtue of benevolence, traditionally understood as love for one’s fellow humans, can be interpreted in different ways. Kant redefines benevolence not as an emotion but as a **determination of the will** — the rational decision to seek the happiness of others. For Kant, to act morally is to act from duty, out of respect for the moral law, not from inclination or affection.

Similarly, utilitarianism, in the spirit of Sidgwick and Bentham, insists on **impartial benevolence**: “the happiness of any one individual is equally important with the equal happiness of any other.” This universalism clashes with the intuition that we owe special obligations to those with whom we share identity or history. Hence arises the question: how can moral impartiality coexist with legitimate partiality — toward family, community, and nation?

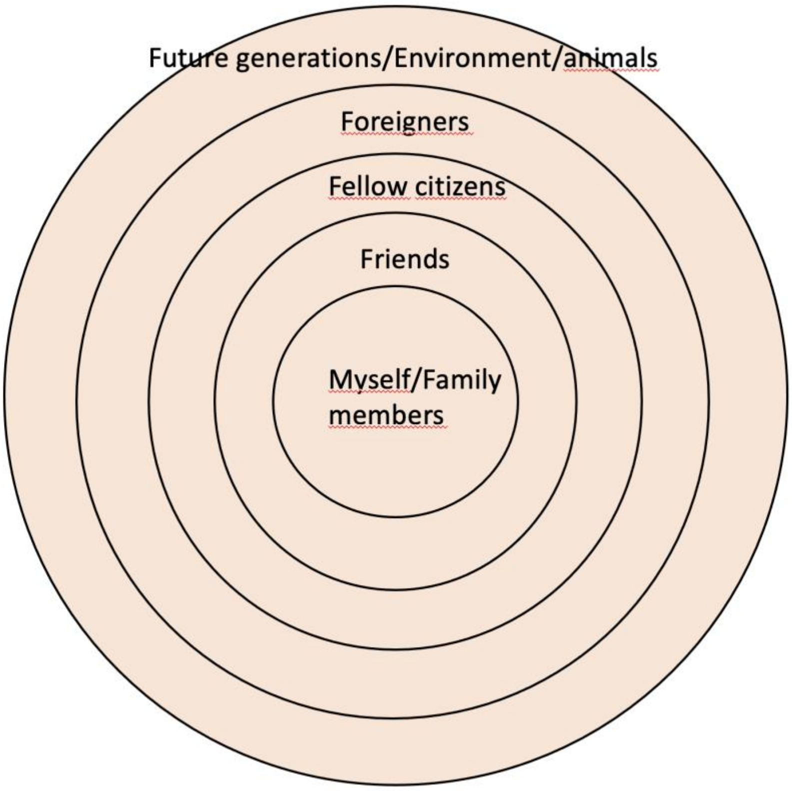


Fig. 1.1 – The expanding circle**.**

Family relationships embody a form of **natural communism**, governed by reciprocity and mutual dependence. Within the family, individuals "give according to their ability and receive according to their need." These intimate relationships justify giving priority to family members, not because they are metaphysically more valuable, but because they are embedded in **social practices of care and trust** that make moral life possible.

Singer would likely agree that such partiality has **instrumental justification** — it promotes flourishing relationships and social stability — but he would deny that these reasons justify ignoring distant suffering. If impartial reasoning shows that one can prevent greater suffering elsewhere without sacrificing something of comparable moral worth, then reason obliges us to expand our concern beyond our kin (See Fig. 1.1 and Box # 3 for the The "Drowing Child Analogy").

At a broader level, membership in a political community also generates **special obligations**. States, as collective moral agents, organize the distribution of goods through taxation and law, aiming at fairness and justice among their citizens. Citizens, in turn, owe each other solidarity because they share in creating and sustaining the institutions that secure collective welfare.

Yet, modern political communities also recognize duties toward non-members — refugees, migrants, and the global poor. The moral legitimacy of the state depends on its commitment to **universal moral principles** such as human rights and humanitarian assistance. In this sense, patriotism — attachment to one's country — must be tempered by **cosmopolitan justice**. As Kant argued in his *Perpetual Peace* (1795), the moral law requires treating every person as an end in themselves, regardless of nationality.

Patriotism, understood as identification rather than mere affection, can play a constructive role in political communities. By fostering a sense of shared belonging, it can strengthen social cohesion, encourage civic virtue, and motivate individuals to contribute to collective goods such as democratic participation, redistribution, and national defense. In this sense, patriotic attachment may be seen as a moral resource: a partial but valuable form of solidarity grounded in common history and institutions.

However, patriotic sentiment is not immune to distortion. When identification with one's nation slips into a belief in its inherent superiority — coupled with the unqualified pursuit of national interests at the expense of others — patriotism degenerates into exclusionary nationalism. This shift marks the dangerous side of patriotic feeling: it narrows the scope of moral concern and permits discrimination, or even violence, against those construed as outsiders or "enemies of the nation." The twentieth century tragically demonstrated how nationalist ideologies, infused with myths of ethnic purity or civilizational destiny, fueled world wars, forced population expulsions, and genocides.

Political theories of nation-building help to explain how this boundary-drawing occurs. As Benedict Anderson famously argued, the nation is an "imagined community," created through shared narratives, language, symbols, and institutions rather than any natural or primordial ties. National identities are therefore shaped and reinforced by state policies. The Italian statesman Massimo D'Azeglio captured this dynamic with his remark after unification: "Fatta l'Italia, bisogna fare gli Italiani" — "We have made Italy; now we must make Italians." Such deliberate construction of national identity can indeed generate solidarity and motivate citizens to view their fates as interdependent.

Yet the very same mechanisms that promote unity can also be used to delineate moral boundaries, drawing a sharp line between "us" and "them." When the imagined community becomes a moral horizon — the primary or exclusive object of loyalty — justice and empathy risk being confined within national borders. The moral circle turns inward. In this way, the constructive potential of patriotism as identification can easily slide into the ethical parochialism of nationalism, presenting a profound challenge: how to cultivate a shared civic identity that supports democratic life without sacrificing universal human concern.

The foregoing discussions reveal that moral life is shaped by a persistent tension between **particular attachments** and **universal principles**. We care deeply about those close to us, yet we also recognize, through reason, that all persons are owed equal moral consideration. To navigate this tension, we require a method capable of systematically bringing our intuitions and our theoretical commitments into coherence.

Rawls's method of **reflective equilibrium** provides precisely such a framework. It asks us to revise both our intuitive judgments (the "morality of common sense") and our theoretical principles (such as Kantian universality or utilitarian impartiality) until they mutually support one another. Through this iterative process, we can distinguish **legitimate forms of partiality** — such as familial care or civic solidarity — from **illegitimate forms**, such as ethnic bias or exclusionary nationalism.

Reflective equilibrium does not demand that we abandon our special relationships. Rather, it requires us to integrate them within a framework of impartial respect for all persons. Singer's expanding circle can be understood as the expression of this rational integration: a moral outlook in which particular loyalties are recognized, valued, and situated within an ever-widening horizon of concern.

The aim is not to erase local attachments, but to prevent them from confining the scope of our moral imagination. Moral progress, on this view, lies in expanding — not contracting — the circle of moral concern, guided by empathy, reflection, and the universal reach of reason.

Our challenge is to cultivate a morality that honors the depth of our particular attachments without allowing them to confine our moral imagination. True moral progress, as Singer suggests, lies in expanding — not abolishing — the circle of moral concern, guided by reflection, empathy, and the universal reach of reason.

Box # 1.3: The “Drowing Child Analogy”

In *“Famine, Affluence, and Morality”* (1972), **Peter Singer** presents one of the most influential moral arguments of the 20th century. His essay challenges our ordinary moral intuitions about **charity**, **duty**, and **moral distance**, arguing that affluent people in wealthy societies have a **strong moral obligation** — not just a moral option — to prevent suffering and death from poverty and famine when they can do so *without sacrificing something of comparable moral importance*.

Singer wrote the essay in response to the **Bengal famine (1971)**, in which millions of people in East Pakistan (now Bangladesh) were suffering from starvation and displacement.
He uses this case to expose what he calls **“a radical revision of our moral conception of duty and charity.”**

Singer’s argument can be summarized in **three main premises** and a **conclusion**:

Premise 1: Suffering and death from lack of food, shelter, and medical care are bad.

This is uncontroversial: everyone agrees that such suffering is morally bad and should, if possible, be prevented.

Premise 2: If it is in our power to prevent something bad from happening, without thereby sacrificing anything of comparable moral significant, we ought to do it.

This is the **principle of preventing bad occurrences** — the moral core of the essay.

Singer stresses that it is not about doing good *supererogatorily* (beyond duty), but *morally obligatory* action.

Premise 3: It is in our power to prevent suffering and death from lack of food, shelter, and medical care, without thereby sacrificing anything of comparable moral significance.

For most people in affluent societies, giving away some money means giving up luxuries — not basic necessities.

Conclusion: Therefore, we ought morally to give far more than we currently do.

This means that what is usually called *charity* should be seen as *moral duty*.
Failing to give, when one can, is **morally wrong** — comparable to allowing harm to occur when one could easily prevent it

The Famous “Drowning Child” Analogy

In support of Premise 2 Singer presents a vivid thought experiment for illustrating his principle of impartiality and proximity:

“If I am walking past a shallow pond and see a child drowning in it, I ought to wade in and pull the child out. This will mean getting my clothes muddy, but this is insignificant compared to the death of the child.”

If we accept that it would be wrong to let the child drown when we could save them easily, then — Singer argues — there is **no moral difference** between that case and failing to send money to save children dying from famine abroad.
Physical proximity or national boundaries do not change moral responsibility.

Singer argues that **common-sense morality** wrongly distinguishes between:

- *Duty* (what we are morally required to do), and
- *Charity* (what is morally good but not required).

He rejects this distinction, calling it **psychologically comforting but morally indefensible**. In a globally connected world, we have the *same obligation* to help distant strangers as nearby ones

Singer draws on the **utilitarian principle of impartiality** — the idea that *"the interests of every being affected by an action are to be taken into account and given the same weight."*

Thus, from the standpoint of reason:

- My child's suffering and a stranger's child's suffering have *equal moral weight*.
- The only morally relevant question is how to reduce the most suffering overall

Singer's conclusion is particularly demanding. He insists that affluent people should in priciple **give until the point of "marginal utility"** — that is, until giving more would cause *as much suffering to oneself* as it would relieve in others.

In his words:

"We ought to give until we reach the level of marginal utility — that is, the level at which, by giving more, I would cause as much suffering to myself or my dependents as I would relieve through my gift." Singer, however, offer a more moderate formulation in addition to the maximalist one: **"We should prevent bad occurrences unless, to do so, we had to sacrifice something morally significant."** What he exactly means for "morally significant" is open to interpretation.

This means that **our current consumer lifestyle is morally indefensible** as long as preventable suffering exists elsewhere.

Singer's essay effectively:

- Collapses the distinction between *charity* and *duty*;
- Argues for a **global moral community** (consistent with cosmopolitanism);
- Rejects *moral distance* and *partiality*;
- Extends utilitarian reasoning into **practical global ethics** — setting the foundation for the modern **Effective Altruism movement**.

In short, Singer's argument in "Famine, Affluence, and Morality" is that distance, nationality, and social connection make no moral difference to our obligation to prevent suffering — and that failing to aid when one can do so easily is morally wrong, not merely less good.

1.6 Identity and Roots: reconciling Universal Obligation and the politics of Recognition

The preceding discussion highlighted a central tension in moral life: while moral reason urges us toward impartial concern for all persons, our capacities for care and responsibility are formed within particular relationships and communities. Universal obligation, as articulated by Kantian, utilitarian, and cosmopolitan theories, cannot operate in abstraction from the social contexts in which moral agents are embedded. Partial attachments are not merely obstacles to moral universality; they are among the conditions that make moral motivation and agency possible.
This raises a deeper question: how should identity, belonging, and rootedness be understood within a moral framework that affirms equal concern for all?

Thinkers such as Immanuel Kant, Simone Weil, Peter Singer, and Thomas Pogge articulate cosmopolitan visions of universal moral obligation that transcend borders, preferences, and identities. Meanwhile, thinkers such as Charles Taylor, Kwame Anthony Appiah, Francis Fukuyama, and communitarians emphasize that human beings do not exist as abstract moral atoms but as situated selves, whose identities depend on recognition, narrative, and cultural rootedness.

This section explores how these two perspectives—the ethics of universality and the politics of recognition—can be understood not as opposites but as mutually enriching. It argues that universal moral obligation must be realized through the particular identities that shape human life, and that rootedness, when properly understood, is not an obstacle to universality but its condition.

Sen and the "illusion of a singular identity"

Sen's critique of the "illusion of singular identity" in *Identity and Violence* (2006) is especially important here, for it directly challenges the idea that rootedness and cosmopolitan obligation stand in inherent tension. Sen argues that the real danger lies not in identity as such, but in its reduction to a single, allegedly destiny-defining category—an intellectual and political simplification at the heart of sectarianism, ethnic nationalism, and civilizational conflict. The claim that one is *nothing but* a member of one group (a religion, a nation, an ethnicity, a "civilization") erases the multiple affiliations that constitute a human life and exposes individuals to manipulation by movements claiming to defend their "true" identity. Against this fatalistic essentialism, Sen emphasizes that each person is a plurality: a speaker of a language, bearer of a profession, participant in a culture, member of a family, citizen of a polity, moral agent, and—crucially—a human being among humans. Individuals must reason about their affiliations, prioritizing them according to context and moral reflection. Once we recognize this plurality, the supposed opposition between rootedness and universal obligation begins to soften: I can be rooted in a local culture and committed to universal moral concerns because I already inhabit many identities simultaneously, among which belonging to humanity is one. The plurality of identity becomes the bridge allowing rootedness to enrich rather than undermine cosmopolitan ethics.

Simone Weil: The Need for Roots and Universal Obligation

In *The Need for Roots* (1949), Simone Weil presents a moral philosophy grounded in obligation rather than rights. For Weil (1949/2002), "a right is not effective by itself, but only by virtue of the obligation to which it corresponds" (p. 3). Rights, she contends, are derivative; what is fundamental is the obligation each person owes to others simply by virtue of their humanity. Weil's concept of obligation is universal and absolute. It does not depend on reciprocity, citizenship, or proximity. Every person, regardless of their condition, is the object of an unqualified moral duty.

At the same time, Weil insists that human beings have a need for roots, "one of the most important and least recognized needs of the human soul" (p. 41). Rootedness—in family, culture, tradition, and place—is necessary for moral and psychological flourishing. The uprooted person, deprived of connection to a community and history, suffers a profound loss of meaning. Yet Weil warns against the perversion of rootedness into exclusionary nationalism or collective egoism. True rootedness, she writes, "is participation in the life of a community which preserves certain particular treasures of the past and certain particular expectations for the future" (Weil, 1949/2002, p. 44). In this sense, she anticipates later debates on recognition: identity requires being seen and valued by others, but this need for recognition must remain anchored in the primacy of moral obligation to all humanity.

Here a link with Sen becomes clear. Weil fears the moral vacuum created by uprootedness; Sen fears the violence unleashed when one root is absolutized at the expense of all others. Weil diagnoses the moral pathologies produced by the loss of roots; Sen diagnoses the pathologies produced when roots are absolutized. Both reject the idea that identity is destiny. For Weil, obligation precedes all social categories; for Sen, no single category may claim to exhaust who a person is.

Charles Taylor: From Honor to Dignity

Charles Taylor, in *The Politics of Recognition* (1992), offers a powerful account of how modernity transformed the moral grammar of identity. In premodern hierarchical societies, recognition took the form of honor, reserved for the elite. Modern egalitarian societies replaced this with the notion of inherent dignity—the idea that all humans are worthy of respect simply as persons. However, Taylor argues that the modern demand for equality has developed into a demand for the recognition of particular identities. Individuals and groups seek acknowledgment not only as human beings but as bearers of distinct cultural, gendered, religious, or ethnic identities.

This is because, as Taylor explains, "our identity is partly shaped by recognition or its absence, often by the misrecognition of others, and so a person or group of people can suffer real damage, real distortion, if the people or society around them mirror back to them a confining or demeaning picture of themselves" (Taylor, 1992, p. 25). Recognition thus becomes a vital human need, not a matter of mere courtesy or tolerance. Without it, individuals internalize a sense of inferiority and lose the capacity for authentic self-realization. The politics of recognition aims to correct historical and cultural patterns of misrecognition that undermine dignity.

Yet this development poses a challenge to universalist moral theories. If everyone's identity is particular, culturally situated, and partially incommensurable, can there still be a common moral law binding all? Taylor insists that the politics of recognition must operate within a shared framework of universal respect—what he calls "a fusion of horizons"—where dialogue among diverse identities takes place on the basis of mutual dignity (p. 67). In this sense, Taylor's project extends rather than opposes Weil's: both affirm that the human need for rootedness and identity is real, but that it must remain open to a universal dimension of obligation.

Sen's concern overlaps here. He accepts that recognition of particular identities is indispensable, but he warns that when any one identity—religion, nation, or "civilization"—is treated as uniquely fundamental, it becomes a tool of division and violence. Taylor thus underscores the need for recognition; Sen shows that recognition must itself recognize plurality.

Appiah: The Ethics of Identity and Rooted Cosmopolitanism

Kwame Anthony Appiah, in *The Ethics of Identity* (2005), builds upon Taylor's insight but reframes it within a cosmopolitan ethical perspective. For Appiah, identities are ethically significant because they shape the narratives through which individuals understand themselves and their responsibilities. Yet identities are also socially constructed, contingent, and multiple. To identify too strongly with a single group or culture is to risk moral myopia and the kind of "illusory singularity" Sen criticizes.

Appiah proposes what he calls rooted cosmopolitanism—the idea that we can be attached to our local traditions and communities while recognizing our moral obligations to all human beings. "Cosmopolitanism," he writes, "begins with the simple idea that we are citizens of the world. But it doesn't mean that we have to give up our local identities, our families, our nations, or our cultures" (Appiah, 2005, p. 222). Moral life, in this view, is lived through particular identities, not in abstraction. Our universal duties are exercised through concrete relationships, practices, and institutions.

In this way, Appiah reconciles the universalist ethics of Kant, Singer, and Pogge with the particularism of Taylor and Weil, and he is closely aligned with Sen's plural humanism. The cosmopolitan, for Appiah, is not a placeless citizen of nowhere but a person who understands that local attachments and global responsibility are complementary rather than mutually exclusive.

Fukuyama: Identity, Dignity, and the Politics of Resentment

Francis Fukuyama, in *Identity: The Demand for Dignity and the Politics of Resentment* (2018), situates the struggle for recognition within contemporary political dynamics. He argues that modern democratic societies, having achieved relative material prosperity, increasingly revolve around identity conflicts rather than class-based ones. Building on his earlier concept of *thymos*—the human desire for recognition—Fukuyama claims that both progressive and populist movements are fueled by perceived denials of dignity. “People do not just want material goods,” he writes; “they want recognition of their worth” (Fukuyama, 2018, p. xv).

On the left, movements for gender, racial, and sexual equality seek acknowledgment of marginalized identities. On the right, nationalist and populist movements express resentment at the perceived neglect or humiliation of cultural or national identity. The result is a fragmentation of the public sphere into competing claims for recognition. In this landscape, Sen’s analysis of how singular identities are constructed and weaponized becomes highly relevant: narratives that define “the people” against “outsiders” or “enemies” simplify complex plural identities into stark oppositions.

Fukuyama’s solution echoes Taylor and Appiah: societies must cultivate an inclusive civic identity rooted in democratic values rather than ethnic, religious, or cultural exclusivity. Only such a shared identity can reconcile universal principles of justice with the human need for belonging. Here again, we can see the outline of a rooted cosmopolitanism at the political level: a framework that allows for particular attachments while upholding universal norms.

Cosmopolitan Universalism: Kant, Singer, Pogge (and Sen)

Against this backdrop of identity and recognition, Kant, Singer, and Pogge represent some of the strongest formulations of cosmopolitan universal moral obligation. For Immanuel Kant, morality is grounded in reason and autonomy, not in social recognition or emotional attachment. The categorical imperative requires that we act only according to maxims that could be willed as universal laws, and that we treat every person as an end in themselves (Kant, 1785/2012). Kant’s ethics is formally universal: it abstracts from all empirical conditions—race, nationality, culture, or preference—to derive a law binding on all rational beings. This impartiality is both the strength and the challenge of Kantian ethics: it affirms human equality but risks detachment from lived identity.

Peter Singer, from a utilitarian perspective, radicalizes a similar moral intuition. In “Famine, Affluence, and Morality” (1972), he argues that if we can prevent suffering without sacrificing something of comparable moral importance, we are morally obligated to do so. The moral relevance of distance, nationality, or identity disappears. “If it is in our power to prevent something bad from happening, without thereby sacrificing anything of comparable moral importance, we ought, morally, to do it” (Singer, 1972, p. 231). For Singer, this logic extends to global poverty, animal suffering, and future generations. The moral circle should expand continually, encompassing all sentient beings.

Thomas Pogge reinforces this cosmopolitan perspective by challenging the communitarian claim that moral obligations arise primarily within bounded communities of shared identity. For Pogge, such identity-based limits to justice are both empirically untenable in a globalized world and morally arbitrary. Contemporary economic and political structures bind individuals together across borders, shaping life chances in ways no national community can contain. Because these transnational institutions often disadvantage the global poor, our primary moral duties arise not from identity or cultural membership but from the **negative obligation not to sustain or benefit from unjust global arrangements**. Even if special responsibilities to compatriots exist, they cannot outweigh duties to those harmed by the global rules we participate in. Identity, therefore, cannot determine the scope of justice. Pogge's "moral individualism"—the view that individuals, not communities, are the fundamental bearers of moral claims—aligns closely with Sen's critique of singular identity: both insist that the worth of persons does not depend on the groups they belong to. In this way, Pogge deepens the argument that universal obligation is not a distant abstraction but a concrete demand emerging from the very structures of interdependence that traverse and ultimately relativize our particular identities.

Together, Kant, Singer, and Pogge articulate a powerful cosmopolitan ethic: all persons matter equally, regardless of their political membership or cultural identity, and existing transnational structures must be evaluated according to their impact on this equal moral standing. Yet critics such as Appiah and Taylor worry that such theories can overlook the motivational and cultural sources that make moral action possible. Human beings do not act purely out of reason; they act out of empathy, identification, and recognition—emotions that presuppose rootedness and identity. Sen, for his part, accepts the universalist thrust of cosmopolitanism but insists that it must respect the plurality of our identities and the agency individuals have in interpreting and ordering them.

The contrast between cosmopolitan universalism and the politics of recognition marks one of the central tensions of modern moral thought. Universalist theories affirm the equal moral worth of all persons but risk abstraction from lived attachments, while recognition-based approaches restore the importance of identity and belonging but risk moral fragmentation. **Simone Weil** insists on universal obligation while defending the necessity of roots; **Immanuel Kant**, **Peter Singer**, and **Thomas Pogge** articulate demanding forms of moral impartiality; **Charles Taylor**, **Kwame Anthony Appiah**, and **Francis Fukuyama** emphasize recognition and identity; and **Amartya Sen** exposes the dangers of absolutized, singular identities while defending a plural, reasoned self-understanding.

The synthesis that emerges can be described as *rooted cosmopolitanism*: a view in which universal moral obligation is not opposed to identity but realized through it. Moral concern becomes effective only when anchored in concrete relationships, practices, and institutions, while particular identities remain open to universal claims of justice. Recognition and universality are therefore not rival principles but mutually dependent dimensions of moral life. The challenge for moral and political philosophy is not to choose between them, but to show how universal equality and particular belonging sustain one another in a shared ethical horizon. The task, therefore, is to integrate these perspectives: to affirm the universal moral worth of every person while recognizing that this worth is lived and experienced through particular cultures, relationships, and communities.

1.7 The Expanding Scope of Negative Duties, and the Intergenerational Justice

In the modern world, even negative duties — those that require only non-interference — have become increasingly demanding. It is no longer sufficient to avoid direct harm; one must also refrain from **indirectly contributing to harm through systemic participation**.

(a) Environmental Harm

Our carbon-intensive lifestyles illustrate this transformation. Activities such as driving, flying, or consuming energy are not morally neutral. They contribute to global warming, which harms vulnerable populations and threatens the survival of future generations. Therefore, actions like reducing energy consumption, eating locally, and limiting emissions express the **negative duty not to harm others**. Failing to act is not mere omission but complicity in collective wrongdoing.

Henry Shue (1980) argues that when harmful institutional practices systematically benefit some while imposing serious risks on others, those who have benefited most bear primary responsibility to prevent and remedy the harm. In the context of climate change, this reframes mitigation and adaptation not as acts of charity, but as the fulfillment of negative duties not to deprive others of the basic conditions of security and subsistence.

(b) Consumer Responsibility

The same applies to **consumer behavior**. Purchasing cheap products made in exploitative conditions perpetuates harm. Thomas Pogge (2002) describes this as "institutional harm": affluent societies sustain global systems that systematically disadvantage the poor. Our moral responsibility thus extends beyond personal virtue to participation in **structures of injustice**. The duty "not to harm" entails not benefiting from the suffering of others.

(c) Governmental and Structural Harm

State policies also have global ethical implications. Tariffs and agricultural subsidies in wealthy countries can devastate farmers in developing economies. When such policies entrench inequality, they violate negative duties by perpetuating foreseeable harm. Justice, therefore, demands not only fair distribution within nations but also **institutional reform at the global level**.

As harms become more complex and diffused, negative duties increasingly require **positive institutional enforcement**. Protecting people from harm now involves preventing environmental degradation, regulating markets, and ensuring fair trade. Governments must act affirmatively to safeguard even the most basic negative rights.

This realization confirms one of the central claim of this essay: **duties precede rights**. The effective protection of rights — whether to life, liberty, or livelihood — depends on the fulfillment of active duties by both individuals and institutions. The moral landscape has expanded from personal virtue to **collective responsibility**.

Duties to Future Generations

One of the most profound implications of this moral expansion concerns **future generations**. Why should we sacrifice for people who do not yet exist? What have they ever done for us? At first glance, such sacrifice appears to be a **one-way street** — lacking the reciprocity or mutual advantage that typically motivates moral obligation.

Yet humankind now possesses an unprecedented capacity to shape the future, for better or worse. We can destroy the conditions for human life or preserve them. Whatever their cultural differences, future people will need cultivable land, clean water, and a stable climate. Doing nothing — continuing to pollute, exhaust, and destroy — cannot be morally acceptable.

Philosophers such as John Rawls and Derek Parfit have provided conceptual frameworks for these obligations. In Rawls's "original position," rational choosers design principles of justice without knowing their place in society. If they also do not know **which generation** they belong to, they would rationally ensure that each generation inherits sufficient resources to sustain a good life. Justice between generations, for Rawls, requires that each passes on to the next "at least as much as it received."

Derek Parfit argues in *Reasons and Persons* (1984) that future generations matter morally just as much as present people, even though they do not yet exist. Because many of our choices determine *which* people will exist, we cannot rely on ordinary person-affecting ideas of harm; instead, we must use **impersonal principles** that judge actions by their overall effects on future well-being. From this, Parfit concludes that we have strong duties to avoid causing predictable future suffering, to ensure that future lives go well, and to design institutions that protect the long-term future. Our obligations therefore extend beyond preventing harm to particular individuals: they include creating a world in which future people can live good lives (see Box # 4).

Several arguments reinforce this duty:

1. **Overlapping Generations:** Care for current people naturally extends to care for their descendants. Concern for one's children implies concern for the children of their children.
2. **Common Stewardship:** Present generations have no right to exhaust or poison resources they did not create. We are stewards of the earth's life-sustaining capacities, bound to pass them on intact.

3. **Inheritance of Public Goods:** Each generation inherits not only natural but also cultural and institutional goods — art, knowledge, infrastructure. Justice requires that we preserve and renew these legacies.
4. **Self-Transcendence:** Concern for future generations also fulfills a deep human need for **self-transcendence** — the ability to find meaning beyond one's lifespan. By acting for the future, individuals overcome the limits of mortality and participate in a continuity larger than themselves. In this sense, caring for the future enriches the present; it is both altruistic and self-fulfilling.

Thus, the **circle of moral concern**, to borrow Peter Singer's term, must expand — from family to nation, from humanity to future humanity.

Box # 1.4: Why We Should Care About Future Generations — Derek Parfit's Argument

In *Reasons and Persons* (1984), Derek Parfit offers one of the most influential accounts of why we owe moral duties to future generations. His starting point is simple: the value of a person's life does not depend on when that person exists. If our actions today will predictably affect the quality of future lives, then those lives matter morally.

Yet Parfit shows that this apparently intuitive idea leads to a deep philosophical puzzle. Many of our long-term decisions—such as adopting a climate policy, building new infrastructure, or regulating pollution—do not merely affect future people; they also determine *which* people will come into existence. Slight differences in policy, timing, or economic conditions alter who meets whom, who has a child when, and therefore which individuals are eventually born.

This insight generates the **Non-Identity Problem**. Suppose a government adopts Policy A, which leads to moderate pollution. Because this policy alters future economic conditions, it also affects when people conceive children. Decades later, a particular child—call her Anna—is born with a mild respiratory condition caused by the polluted environment. If the government had adopted the cleaner Policy B, Anna would *not* have existed at all; different conceptions would have occurred, producing different people. Can Anna claim that Policy A wronged her? Parfit argues she cannot. Although her life is somewhat worse than it might have been under a cleaner policy, the "better outcome" for Anna would have been **never to exist**, which is not a real alternative for her. In this sense, Policy A harms *no particular person*, even though it makes the world worse.

Parfit's conclusion is striking. Traditional moral theories, which explain wrongdoing in terms of harming specific individuals, cannot account for why we judge Policy A morally objectionable. To make sense of our intuitions, we must accept an **impersonal moral principle**: some actions are worse *not because they wrong particular people*, but because they lead to a world with less well-being overall. On this view, we should care for future generations not simply to avoid violating rights, but because the quality of future lives is something we ought, in general, to promote.

Parfit's analysis thus reframes our responsibilities toward the future. Even though future people do not yet exist—and even though our choices determine their very identities—we have strong moral reasons to avoid leaving them a damaged planet, depleted resources, or avoidable suffering. What ultimately matters is not who comes to exist, but **how good their lives can be**. This impersonal perspective provides a powerful justification for taking long-term challenges such as climate change, biodiversity loss, and technological risk seriously.

1.8 Duties as the Foundation of Rights: From Moral Obbligation to Global Justice

The Institutional Priority of Duties

Human rights discourse often begins with entitlements: rights to liberty, security, subsistence, or political participation. These rights are treated as morally fundamental claims grounded in equal human worth. Yet rights do not secure themselves. Their practical significance depends entirely on the duties that correspond to them and the institutional arrangements through which those duties are specified and enforced.

To say that duties have priority over rights is therefore not to diminish rights, but to clarify their structure. A right is meaningful only if someone is obligated to act—or to refrain from acting—in ways that protect the interest the right secures. Where no duties are clearly assigned, rights remain moral aspirations rather than enforceable claims.

Moreover, the duties corresponding to a right are rarely singular or simple. Protecting even the most familiar rights requires a complex network of responsibilities. To secure freedom from violence, for example, it is not enough that others refrain from assault; there must also be institutions capable of preventing threats, deterring wrongdoing, and responding when violations occur. Likewise, ensuring access to essential goods requires more than isolated acts of charity; it demands sustained structures that enable provision, coordination, and accountability.

This reveals an important structural feature of rights: each right generates multiple layers of obligation. At minimum, there are duties not to violate it directly. But there are also duties to shield individuals from third-party interference and duties to assist when protection fails or deprivation occurs. These obligations differ in form and bearer, yet all are necessary if a right is to function as a genuine guarantee rather than a rhetorical commitment.

Modern political institutions transform these layered moral responsibilities into public systems of justice. Through law, taxation, regulation, and enforcement, they allocate burdens, identify duty-bearers, and establish mechanisms of compliance. Rights become effective not through moral recognition alone, but through organized schemes capable of preventing deprivation, protecting against harm, and providing assistance when needed.

Understanding this institutional structure is essential when turning to those rights that are most fundamental. If some rights are preconditions for the enjoyment of all others, then the duties they generate must be correspondingly robust and systematically organized. It is precisely this insight that Henry Shue develops in his account of basic rights.

Basic rights and Institutional Duties: Henry's Shue Framework

It is precisely this question that Henry Shue addresses in Basic Rights. Shue (1980) focuses not on rights in general, but on those rights that are necessary for the enjoyment of all others. Shue identifies two such rights: security rights and subsistence rights.

Security rights are often described as “negative” rights. They include the right to physical security: not to be subjected to murder, torture, assault, rape, or arbitrary violence.

Subsistence rights are typically labeled “positive” rights. They include access to unpolluted air and water, adequate food, clothing, shelter, and the minimal economic conditions required for a healthy and active life.

For Shue, guaranteed security and guaranteed subsistence are simultaneous necessities. Without them, no other rights can be meaningfully exercised. Freedom of peaceful assembly, for example, is conceptually distinct from security from physical assault, but it cannot be exercised without it. Security and subsistence are therefore not optional additions to liberty; they are its material preconditions.

Crucially, both types of basic rights require positive institutional action. It is misleading to think that security rights are satisfied merely by others refraining from harm. The protection of physical security requires police forces, criminal courts, prisons, trained personnel, legal procedures, and systems of enforcement. These institutions must be created, maintained, and funded through taxation and public debt. Physical security is therefore not merely a matter of negative non-interference; it is the product of organized collective action.

The same is true of subsistence rights. Their fulfillment involves more than distributing goods or money to the needy. It requires creating social and economic structures that enable individuals to support themselves: access to employment, education, healthcare, and social protection. Here again, rights depend on sustained institutional duties rather than isolated acts of goodwill.

Shue’s central insight is that basic rights always impose three types of duties: **duties to avoid depriving (negative duties), duties to protect from deprivation, and duties to aid the deprived**. Even rights traditionally classified as negative require all three. The distinction between negative and positive rights therefore obscures the deeper reality: all basic rights depend on active duties.

Among these, the duty to avoid depriving occupies a special place. It is a negative duty: a duty not to cause or contribute to harm. For Shue, this duty is stringent and primary. It is more demanding than duties of assistance because its violation constitutes wrongdoing. Unlike charitable assistance, it does not depend on goodwill or generosity; it prohibits conduct that foreseeably undermines others’ basic rights.

This distinction becomes especially important in the context of climate change. Shue argues that responsibility for preventing harm does not fall equally on all actors. Two principles are central. First, those who cause or significantly contribute to harm bear special responsibility to stop it. Second—and equally important—those who benefit most from harmful institutional arrangements have primary responsibility to correct them. When an activity produces systemic harm, generates concentrated benefits for some, and imposes diffuse costs on others, justice requires that the primary burden of prevention fall on those who have benefited most, who possess greater capacity to respond, and who have helped sustain the harmful structure.

Applied to climate change, this reasoning has powerful implications. The global climate crisis results from long-term fossil fuel combustion, with cumulative emissions concentrated in industrialized states. At the same time, the gravest vulnerabilities are often found in poorer regions that contributed least to the problem. Under Shue's framework, high-emitting societies have benefited most from fossil-fuel-driven development, have contributed most to historical emissions, and possess greater economic and technological capacity to mitigate harm. Their obligation to reduce emissions and support adaptation is therefore not an act of generosity. It is the discharge of a negative duty not to deprive others of the basic conditions of security and subsistence. Climate change threatens both: it endangers physical security through extreme weather, displacement, and conflict, and it undermines subsistence through impacts on food systems, water access, and habitability. Preventing climate harm thus becomes a matter of protecting basic rights.

This reframing marks a decisive shift from charity to justice. Charity implies voluntary assistance—morally commendable but discretionary. Justice, by contrast, concerns rectifying or preventing harm in which one is implicated. It involves binding obligations rather than optional benevolence. If climate harms are foreseeable, avoidable, and structurally connected to institutionalized economic systems, then mitigation is not generosity but compliance with a stringent negative duty. It is a requirement of justice.

Finally, Shue's reasoning extends beyond national borders. Emissions are transboundary, harms are global, and the institutions sustaining fossil-fuel economies operate internationally. The duty to avoid depriving others of basic rights cannot therefore be territorially confined. Justice tracks the structure of harm, not the boundaries of the state. Where systemic activities threaten security and subsistence across borders, the corresponding duties are global in scope

From Domestic Justice to Global Justice

If basic rights depend on institutional arrangements, the next question concerns scope: which institutions bear responsibility for securing them?

Within the modern state, duties are specified through law, enforced by public authorities, and justified through political processes. In this context, the institutionalization of responsibility largely coincides with national borders. However, Shue's analysis does not itself confine duties territorially. Duties to avoid, protect, and aid arise wherever agents and institutional schemes systematically structure individuals' access to the conditions necessary for basic rights.

This suggests a broader principle: **where institutional arrangements foreseeably and avoidably structure individuals' access to the conditions of basic rights, those who participate in, benefit from, or exercise significant influence over those arrangements bear responsibilities proportionate to their role and capacity to effect change**. Justice therefore tracks not mere causal contact, but structured and rule-governed systems of power that shape fundamental interests.

In an interdependent world, domestic institutions operate within global frameworks of trade, finance, intellectual property, environmental regulation, and migration governance. These are not merely diffuse market interactions; they are rule-based systems backed by enforcement mechanisms and asymmetries of bargaining power. In addition, transnational institutions such as the International Monetary Fund, the World Bank, and the World Trade Organization directly shape macroeconomic policy space, debt sustainability, trade rules, and intellectual property protections, including global patent regimes governing access to medicines and technology. Through conditional lending, dispute settlement mechanisms, and regulatory standard-setting, these institutions structure states' fiscal autonomy, development trajectories, and access to essential goods. Where such institutional arrangements systematically influence the realization of security and subsistence, questions of justice arise alongside questions of policy.

This does not imply that all cross-border effects generate justice-based obligations, nor that global institutions are solely responsible for deprivation. Domestic governance failures and local conditions remain morally relevant. However, where international and transnational institutional arrangements foreseeably and avoidably contribute to structural deprivation, responsibility cannot be dismissed as mere humanitarian concern. The issue becomes whether existing institutional schemes are compatible with the duties that basic rights generate.

It is here that contemporary debates on global justice emerge. Thomas Pogge argues that severe global poverty is connected, at least in part, to the design and maintenance of the international economic order. If international institutions foreseeably sustain patterns of avoidable deprivation, then participants in and beneficiaries of those schemes may violate duties not to impose or uphold unjust institutional structures. The wrong, on this view, lies not only in failing to assist, but in contributing to institutional arrangements that undermine basic rights.

Moreover, this account does not require the presence of centralized coercive authority as a condition for justice to arise. Justice is not triggered only where individuals are jointly subject to a sovereign power, but wherever institutionalized and rule-governed arrangements systematically and foreseeably shape access to the conditions necessary for basic rights. Mere causal interdependence or diffuse market effects are insufficient. What matters is whether structured frameworks—such as trade regimes, financial rules, debt conditionality programs, or patent systems—create patterned constraints on security and subsistence that are reasonably avoidable through alternative institutional design. Where such arrangements exist, the absence of formal political authorship or centralized coercion does not dissolve responsibility; it instead raises the question of how institutional power over basic rights is justified and distributed.

By contrast, Thomas Nagel in *The Problem of Global Justice* maintains that stringent principles of justice apply only within sovereign political communities characterized by shared coercive authority and collective authorship. Outside such frameworks, he argues, obligations remain humanitarian rather than justice-based.

Nagel's position rests on a distinctive account of the political conditions of justice. For him, demanding principles of distributive justice arise only where individuals stand in a relation of shared political subjection to a sovereign authority that acts in their name. Because citizens are both authors and subjects of coercive law, they owe one another justifiable terms of cooperation within that institutional framework. In the absence of such collective political authorship—at the global level—individuals and states may owe humanitarian concern, but not duties of justice grounded in shared governance. Justice, on this view, is inseparable from the structure of sovereign political authority.

The disagreement ultimately concerns whether justice tracks sovereignty or institutional power. If justice requires shared political authorship, its scope remains state-bound. If, however, justice arises wherever rule-governed and institutionally structured systems—domestic or transnational—shape individuals' access to the conditions of basic rights, then the absence of global democratic authorship does not eliminate justice claims; it may instead reveal a deficit in existing institutional arrangements.

Understood in this way, the move from domestic to global justice is not an unwarranted expansion but a structural implication of the institutional dependence of basic rights. Wherever institutional schemes—state-based, international, or transnational—systematically and avoidably affect access to security and subsistence, corresponding duties arise in proportion to participation, influence, and capacity for reform. In a world structured by interconnected systems of governance, the boundaries of justice cannot simply be presumed to coincide with the boundaries of the state.

Box # 1.5: Identity Beyond the Self: Parfit and the Dalai Lama on the Ethics of Personhood and Belonging

The question of identity—what it means to be a person, to belong, and to relate to others—lies at the crossroads of metaphysics, ethics, and political life. Across vastly different traditions, the writings of Derek Parfit and the Dalai Lama converge on a strikingly similar critique of the idea that identity is singular, stable, or self-contained. Through the distinct languages of analytic philosophy and Buddhist thought, both challenge the assumption that there exists an indivisible, persisting self that grounds our moral and social existence. In their different registers, each proposes a conception of identity as relational, fluid, and ethically expansive, thereby undermining the conceptual foundations of egoism, sectarian attachment, and nationalist exclusivity.

Parfit and the Dissolution of the Self

In *Reasons and Persons* (1984), Derek Parfit advances one of the most radical challenges to traditional conceptions of personal identity. Rejecting the idea of a single, enduring self that persists through time, Parfit argues that personal identity is "not what matters" in survival. What persists is not a metaphysical subject but a bundle of psychological relations—memory, intention, consciousness, and causal continuity. There is no deeper fact behind these relations; the supposedly permanent self is an illusion imposed upon a series of connected mental events.

Parfit's central insight is that recognizing this truth can be morally transformative. When we cease to see ourselves as metaphysically separate beings, the boundary between "self" and "others" loses its apparent depth. As he famously puts it, believing in a permanent ego makes one feel "imprisoned in a glass tunnel." Seeing through the illusion of selfhood, "the walls of the tunnel disappear," and concern for others expands naturally (Parfit, 1984, pp. 281–283).

The dissolution of the self thus becomes the basis for a broader ethical horizon. Without a rigid metaphysical boundary, egoistic concern is weakened, and altruism becomes rationally and emotionally compelling. Parfit's reductionism yields an ethics of impartiality and compassion—not as a spiritual discipline, but as the natural conclusion of a metaphysically modest account of personhood. If the self is not a discrete entity but a continuity within a larger network of beings, then moral concern can extend outward without contradiction.

The Dalai Lama and the Buddhist No-Self

Within a very different intellectual and spiritual tradition, the Dalai Lama articulates a parallel vision grounded in the Buddhist doctrine of *anatta* (no-self). According to this view, the "I" we ordinarily take as real has no independent or intrinsic existence. What we call the self is a convenient designation for the five aggregates—body, sensation, perception, volition, and consciousness—none of which provides a stable metaphysical core. As the Dalai Lama repeatedly emphasizes, "We are not independent entities; our existence depends on countless causes and conditions" (*The Art of Happiness*, 1998).

This metaphysical insight leads directly to a profound ethical one. The belief in a solid, separate self is the root of egocentric attachment and suffering; the realization of interdependence gives rise to compassion (*karuṇā*) and non-harming (*ahiṃsā*). Because all beings are interconnected and share the wish for happiness, the dissolution of the ego expands one's sphere of moral concern. Compassion becomes not simply a virtue but the most rational response to the nature of reality.

The Dalai Lama extends this logic from the personal to the collective level. He distinguishes between healthy forms of cultural identity and the dangers of nationalism, which he describes as "a kind of collective selfishness" (*Ethics for the New Millennium*, 1999). Just as individuals cling to the illusion of a permanent self, groups and nations cling to rigid identities that generate conflict. Seeing through the illusion of separateness—whether individual or collective—offers a path toward peace and solidarity.

Convergence Across Traditions

Despite the distance between analytic reductionism and Buddhist philosophy, Parfit and the Dalai Lama converge on several foundational insights.

First, both reject essentialism. Parfit denies a metaphysical ego; the Dalai Lama denies an intrinsic self. In both cases, the self is revealed as composite, contingent, and relational.

Second, both replace fixity with continuity and interdependence. For Parfit, what matters is psychological connectedness; for the Dalai Lama, the web of causes and conditions that constitute all beings.

Third, both derive expansive ethical conclusions from the dissolution of the self. Parfit's analytic argument undermines egoism and supports impartial concern for others. The Dalai Lama's spiritual teaching roots compassion and non-violence in the nature of reality.

Their convergence is particularly evident in their moral psychology. Parfit's claim that "the separateness of persons is not a deep fact" resonates with the Dalai Lama's teaching that clinging to the illusion of self is the source of suffering. Both argue that transcending this illusion—through philosophical reflection or spiritual practice—opens the way to greater altruism, empathy, and global responsibility.

Yet their approaches remain distinct. Parfit develops his view through thought experiments and conceptual analysis; the Dalai Lama through contemplative practice and ethical cultivation. Parfit dissolves the ego for philosophical clarity; the Dalai Lama for spiritual liberation and compassion. Together, however, they offer complementary paths toward a view of identity that is non-exclusive, dynamic, and ethically generative.

Identity as Responsibility

In their different idioms, Parfit and the Dalai Lama articulate a vision of identity that moves beyond the isolated individual toward a model of interconnected personhood. Parfit dismantles the metaphysical ego; the Dalai Lama transforms this dissolution into a practice of compassion. Against the modern pathologies of egoism, sectarianism, and nationalism, both affirm that peace and justice begin with a rethinking of who we are. To recognize that we are not fixed, self-sufficient entities but participants in a larger web of relations is to replace the illusion of separateness with the reality of responsibility—an ethical foundation for coexistence in an interdependent world.

Appendix A - From Aristotle to Rawls and Sen: The Evolution of Corrective and Distributive Justice

The distinction between **corrective** and **distributive justice**, first articulated by **Aristotle** in the *Nicomachean Ethics* (Book V), has been one of the most enduring and fruitful divisions in the history of moral, legal, and political philosophy. What began as a reflection on the different ways in which justice manifests itself in the polis has evolved into a conceptual framework for understanding both private and public law, as well as the moral structure of modern societies. From Aristotle's proportional conception of justice to Kant's and Hegel's rational reconstructions, and ultimately to Rawls's and Sen's theories of social fairness and capability, the notion of justice has progressively expanded from an interpersonal virtue to a structural ideal of political and social order.

Aristotle: Justice as Proportion and Equality

In *Nicomachean Ethics* Book V, Aristotle distinguishes between two forms of particular justice — **distributive** and **corrective (or rectificatory)**. Both are instances of justice in the sense of fairness, but they operate in distinct spheres of human relations.

Distributive justice concerns the **allocation of honors, wealth, or other goods** among members of a political community. For Aristotle, the principle governing this distribution is **geometric proportion**: individuals should receive shares that are **proportionate to their merit** or desert as recognized by the community. Justice, in this sense, is not simple equality but proportional equality. Injustice occurs when someone receives more or less than their due share relative to others, upsetting the balance that sustains civic harmony.

Corrective justice, by contrast, governs the sphere of **interpersonal transactions**, whether voluntary (contracts, exchanges) or involuntary (theft, assault, negligence). Here, justice is **arithmetical**, not proportional. When one person wrongs another or gains at another's expense, the judge restores equality by **removing the excess** from one and **adding it to the other**, thereby reestablishing balance. Corrective justice thus has a **retrospective** character: it looks backward to repair a past wrong, whereas distributive justice is **prospective**, concerned with how future benefits and burdens ought to be distributed.

Aristotle's framework reflects the ethical and civic orientation of classical Greek thought. Justice is a virtue of character and a practice of balancing relations within the polis. Yet this virtue-based and teleological conception would later be transformed by the modern shift toward **autonomy**, **rights**, and **rational equality**.

Kant: Corrective Justice and the Reciprocity of Freedom

The **modern turn** in the theory of justice begins with **Immanuel Kant**, whose *Metaphysics of Morals* (1797) redefines justice in terms of **right (Recht)** rather than virtue or desert. For Kant, the fundamental principle of justice is that **each person has an innate right to freedom**, compatible with the equal freedom of all others. Justice is thus not about achieving happiness or proportionality, but about **maintaining the formal conditions of reciprocal freedom** among rational beings.

Within this framework, **private right** — the realm of property, contract, and reparation — embodies what Aristotle called **corrective justice**, but now in a new moral form. Wrongdoing, for Kant, consists in **using another's external freedom as a mere means**, violating the condition of equality that justice requires. The legal remedy restores that equality, not because it maximizes welfare or moral virtue, but because it **reestablishes the rightful relations** between autonomous agents.

Modern legal philosophers such as **Ernest Weinrib** and **Jules Coleman** have developed this Kantian line of thought into a sophisticated theory of **corrective justice** as the moral foundation of private law. According to Weinrib, private law (especially tort and contract law) expresses a **bipolar structure**: it links the doer and sufferer of a wrong in a relational form of equality. The purpose of law, in this view, is not to promote social welfare or efficiency, but to **repair the normative imbalance** caused by the wrongful act. Corrective justice, as Weinrib argues, is "the juridical embodiment of Kantian right."

Thus, in the Kantian tradition, corrective justice ceases to be a function of virtue or proportion and becomes an expression of **reciprocal moral autonomy**. It is justice in the form of **rectification**, where the law upholds the abstract equality of free and rational persons.

Hegel: From Abstract Right to Ethical Life

While Kant conceived justice in the formal terms of equality and freedom, **G.W.F. Hegel** sought to embed these principles within the **concrete institutions** of social life. In the *Philosophy of Right* (1821), Hegel divides the domain of right into three levels: **Abstract Right**, **Morality**, and **Ethical Life (Sittlichkeit)**.

Corrective justice, corresponding roughly to Kant's private right, belongs to the level of **Abstract Right**. Here, individuals are recognized as **bearers of rights —** persons who own property, enter into contracts, and can be wronged by others. When such wrongs occur, law intervenes to **restore the conditions of recognition** and mutual respect. Hegel thus interprets corrective justice as an instance of **reconciliation**: a wrong disturbs the unity of recognition, and legal redress restores it.

However, Hegel moves beyond the formalism of Kantian right. For him, justice cannot remain at the abstract level of individual transactions; it must culminate in **Ethical Life**, where freedom is realized through **institutions —** the family, civil society, and the state. Within this higher sphere, the problem of justice becomes **distributive** and **institutional**: how to ensure that the social and political order embodies and sustains the freedom of all.

Thus, in Hegel's dialectical framework, **corrective** and **distributive** justice are not separate but **integrated moments** of the realization of freedom. Corrective justice pertains to the **micro-level** of interpersonal relations, while distributive justice belongs to the **macro-level** of the ethical and political community.

Rawls: Justice as Fairness and the Structure of Society

The Kantian and Hegelian reorientations paved the way for modern theories of **distributive justice**, particularly that of **John Rawls**. In *A Theory of Justice* (1971), Rawls proposes a conception of "justice as fairness," designed to determine the **basic structure** of a well-ordered society.

Rawls shifts the focus from individual virtue or corrective rectification to the **institutional distribution of rights, duties, and advantages**. Distributive justice, in his framework, concerns how the **"basic goods"** of society — liberties, opportunities, income, wealth, and self-respect — are to be allocated among citizens viewed as **free and equal persons**.

Rawls's famous **two principles of justice — ** the **equal basic liberties principle** and the **difference principle —** express this idea. The first ensures an equal scheme of fundamental rights and freedoms for all; the second allows social and economic inequalities only if they work to the **advantage of the least well-off** and are attached to positions open to all under fair equality of opportunity.

Unlike Aristotle's merit-based proportionality, Rawls's conception is grounded in **moral equality** and **fair reciprocity**. It is also distinct from Kantian corrective justice: while the latter concerns rectification between individuals, Rawls's distributive justice concerns the **design of social institutions** that ensure fairness in the long run.

Rawls thus translates the Kantian moral ideal of autonomy into a **political conception** of justice, where fairness is achieved not through individual virtue or compensation but through **just institutions** that structure the basic terms of social cooperation.

Sen and Nussbaum: The Capabilities Approach

While Rawls provided the most influential modern account of distributive justice, later theorists such as **Amartya Sen**and **Martha Nussbaum** criticized his focus on the distribution of **primary goods** and proposed instead the **capabilities approach**.

For Sen, justice should not be measured merely by the distribution of resources or rights, but by what individuals are **actually able to do and to be —** their **capabilities** to lead lives they have reason to value. Nussbaum extends this idea by proposing a list of **central human capabilities** (such as life, health, education, and political participation) that must be secured for all citizens to achieve genuine human flourishing.

This approach reinterprets **distributive justice** in a more substantive and pluralistic way. Where Rawls emphasized **formal equality** and fair procedures, Sen and Nussbaum emphasize **substantive freedom —** the real opportunities individuals possess. Their conception thus integrates **ethical, social, and developmental dimensions** of justice, bridging the gap between moral philosophy and public policy.

In doing so, the capabilities framework reconnects modern distributive justice to the **Aristotelian concern for human flourishing (eudaimonia)**, while retaining the modern emphasis on **freedom and equality** as its normative foundation.

Conclusion: The Continuity and Transformation of Justice

From Aristotle to Sen, the theories of justice have evolved through distinct conceptual phases, yet a common thread persists: the pursuit of **equality —** whether as proportion, reciprocity, recognition, or capability.

- **Aristotle** conceived justice as a virtue of proportion and balance within the polis.
- **Kant** transformed it into a rational principle of reciprocal freedom.
- **Hegel** embedded it in the ethical life of institutions.
- **Rawls** reconstructed it as fairness in the structure of society.
- **Sen and Nussbaum** reimagined it as the substantive capability to live a fully human life.

Corrective justice today remains the moral core of private law — the domain of reparation and responsibility between individuals — while **distributive justice** has become the guiding ideal of political and social order. Their evolution reflects the widening scope of the idea of justice itself: from the equilibrium of exchanges among citizens to the fairness of global institutions and human capabilities.

Ultimately, both forms of justice represent complementary dimensions of the same moral horizon — the commitment to a world in which individuals are recognized as **free and equal agents**, capable of living **just lives within just institutions**.

Appendix B – Kant's Moral Philosophy: Autonomy, Duty, and the Categorical Imperative

Immanuel Kant (1724–1804) develops a radically distinct foundation for morality, one that does not rely on happiness, virtue, or social outcomes but on **rational autonomy** and the **authority of moral duty**. In the *Groundwork of the Metaphysics of Morals* (1785) and the *Critique of Practical Reason* (1788), he argues that moral laws must be grounded in pure practical reason and that every rational agent possesses an intrinsic dignity as a self-legislating being.

Contrasting Ethical Frameworks: A Preliminary Orientation

Kant's theory is best understood against the background of three dominant ethical approaches:

- **Utilitarianism**: judges actions by their tendency to maximize aggregate happiness.
- **Libertarianism**: prioritizes individual liberty and property rights; redistribution is often viewed as coercive.
- **Communitarianism**: conceives justice as promoting virtue or distributing goods according to moral desert.

Kant rejects all three as bases for morality. Consequences, preferences, and social ideals cannot determine what is right; moral principles must hold **unconditionally** and for all rational beings. Moral worth therefore cannot derive from producing good results, advancing self-interest, or realizing a vision of the good life as advocated by communitarians. It arises solely from **acting from duty**.

Freedom as Autonomy

Kant's account of freedom is central to his ethics. Moral agency requires **autonomy**, understood as **self-legislation by reason**. Acting autonomously is not equivalent to acting on one's desires or inclinations; indeed, actions grounded in impulse or social pressure count as **heteronomous**, because they are determined by forces external to one's rational will.

To act freely is to act according to a principle one could rationally will as a law for all. Autonomy therefore gives rise to moral obligation: if the will is the source of the moral law, it is also bound by it.

Duty, Intention, and Moral Worth

For Kant, the moral quality of an action depends entirely on its **maxim**—the principle on which the agent acts—and on whether this maxim is adopted **from duty**. An action that conforms to duty but is motivated by inclination (e.g., generosity performed for praise or pleasure) lacks genuine moral worth. Only action performed *because* it is the right thing to do manifests the proper exercise of rational autonomy.

The Categorical Imperative: The Form of Moral Law

Kant distinguishes **hypothetical imperatives**—rules that apply conditionally, given particular desires—from **categorical imperatives**, which command universally and unconditionally. The categorical imperative expresses the structure of moral obligation itself.

Kant formulates it in several equivalent ways; the two most significant are:

1. **Formula of Universal Law**
   "Act only according to that maxim whereby you can at the same time will that it should become a universal law."
   This test requires asking whether one's maxim could coherently be willed as a law governing all rational agents. Maxims that would undermine the very practice they rely on—such as lying—cannot be universalized and are therefore morally impermissible.
2. **Formula of Humanity**
   "Act so that you treat humanity, whether in your own person or in that of any other, always as an end and never merely as a means."
   Because rational beings are capable of self-legislation, they possess inherent dignity. They must never be used *merely* as instruments for achieving another's ends, even for beneficial purposes.

These two formulations jointly express Kant's core insight: the moral law arises from the rational capacities that all persons share, and therefore binds all rational agents equally.

Respect, Dignity, and the Basis of Human Rights

Kantian ethics grounds moral equality in the **dignity of rational agency**, not in ownership of oneself (as libertarians claim) nor in the capacity to generate utility (as utilitarians claim). Respect for persons follows from their status as autonomous lawmakers within the moral domain. Freedom and moral responsibility are thus inseparable: to be free is to bind oneself to the law that reason prescribes.

This framework provides the philosophical basis for modern conceptions of **human rights**, which claim universal validity independent of cultural norms, political interests, or empirical differences among persons.

Illustrative Applications

- **Keeping promises** satisfies the test of universalizability and expresses respect for others' autonomy.
- **Assisting those in need** reflects a maxim that rational agents could will universally and demonstrates recognition of others' dignity.
- **Lying, coercing, or manipulating others** fails both the universalization test and the requirement to treat persons as ends.

Kantian Ethics as an Alternative Moral Paradigm

Kant's moral philosophy offers a systematic alternative to both consequentialist and libertarian theories. It locates the authority of morality not in outcomes, preferences, or conceptions of virtue but in the self-legislating capacity of rational agents. To act morally is to act from respect for the moral law—a law one gives to oneself through reason, and a law that applies universally to all persons.

In this way, Kant provides a framework for understanding justice, rights, and moral equality that aspires to be valid *beyond* contingent interests and cultural contexts, grounding ethics in the universal structure of rational agency itself.

PART II

Chapter 2

CONSEQUENTIALISM, WELFARISM, AND UTILITARIANISM

2.1 Introduction

No philosophical tradition has shaped modern public policy, economics, and global ethics more than utilitarianism. Even where it is not explicitly invoked, its core intuition — that actions and institutions should be judged by their consequences — still underlies much of contemporary decision-making: from cost–benefit analysis to climate mitigation, from public health triage to humanitarian intervention. The vocabulary of “maximizing welfare”, “efficiency”, and “optimal allocation” is the secular language of global governance — and it is, historically, utilitarian language.

Yet utilitarianism is also one of the most contested and criticised doctrines in the history of ethics. Its reduction of value to welfare, its aggregation across persons, its difficulty in capturing rights, dignity, or fairness, and its famous puzzles around future generations have generated a vast literature of critique — from Kant to Rawls, from Sen to Nussbaum. Many contemporary theorists of global justice define their own positions in opposition to utilitarianism: global egalitarianism rejects simple aggregation; the capability approach rejects welfarism; international human rights theory rejects the subordination of rights to total utility.

But this should not mislead us. Utilitarianism remains indispensable to the philosophical genealogy of global justice. Without utilitarianism, we would not even have the conceptual landscape against which later theories — Rawls’s justice as fairness, Sen’s capability freedom, Pogge’s global institutional critique — define themselves. Its historical development, from Bentham’s hedonistic calculus to Mill’s richer conception of human flourishing and dignity, set the modern terms of the debate. Its methodological innovations — aggregation, marginal analysis, the idea of impartiality extended to all sentient beings — became the foundation on which welfare economics and later global ethics were built. Even the most sophisticated critiques of utilitarianism depend on a deep engagement with its central questions: *what should be maximized, for whom, and why?*

For this reason, a serious philosophy of the global order cannot bypass utilitarianism. It must understand its internal logic, its evolution, its limits, and its enduring relevance. This chapter therefore returns to utilitarianism not in order to revive it in its classical form, but in order to recognise its unique status: the first modern attempt to treat all persons — everywhere — as equal subjects of moral concern, within a single, general theory. Whether one ultimately defends utilitarianism, reforms it, or rejects it, the contemporary conversation about global justice still takes shape in the conceptual horizon it created.

2.2 Consequentialism, Teleology, and the Utilitarian Conception of the Good

Consequentialism affirms that the moral rightness or wrongness of an action depends solely on its **consequences**. What makes an action morally right is not the intention behind it nor its conformity to duty, but rather the **goodness or badness of the outcomes** it produces. Thus, the morally right act is the one that yields the **best overall consequences** among the available alternatives.

Consequentialism is therefore a **maximizing doctrine**: moral agents are not merely encouraged but required to maximize the good. Even when all available options have undesirable results, the right choice remains the one that **minimizes harm** or **produces the least bad outcome**.

John Rawls (1971) captures this logic when he defines a **teleological ethical theory** as one in which "the good is defined independently from the right, and then the right is defined as that which maximizes the good" (p. 24). This distinction between the **right** and the **good** helps delineate the boundary between **teleological** (consequentialist) and **non-teleological** (deontological) ethics.

The fundamental difference between teleological and non-teleological ethics lies in the order of moral evaluation.

- **Teleological ethics** begins by evaluating the goodness of a **state of affairs** and defines the rightness of an act in terms of its contribution to that state.
- **Non-teleological ethics**, by contrast, assigns **intrinsic moral value** to certain actions regardless of their consequences.

For instance, the claim that **breaking a promise is wrong in itself**, even if it produces good results, represents a **non-teleological** or **deontological** standpoint. Conversely, the teleological view holds that the right act is the one that produces the **best state of affairs** overall.

Consequentialism is thus a **specific form of teleological ethics**: it defines the rightness of actions solely in terms of their **goodness of consequences**. To apply it meaningfully, however, one must specify what counts as "good" — that is, provide a **theory of value**.

Consequentialism can be expressed in **formal terms** as a principle of maximization. The **expected moral value** of an action can be computed as the sum of its possible outcomes, weighted by their respective probabilities and values.

For example, if an action A has two possible outcomes — one with probability 0.7 and value 10, and another with probability 0.3 and value –2 — the **expected value** of A is:

$$E(v) = (10\times0.7) + (-2\times0.3) = 6.4$$

According to consequentialist reasoning, if no other available action has a higher expected value, A is the **morally right act**.

However, consequentialism alone does not constitute a **complete ethical theory**. It requires:

1. A **theory of value (the good)** — what outcomes are intrinsically desirable; and
2. A **principle of aggregation** — how individual goods are combined to form an overall measure of moral worth.

These additional components are what transform consequentialism into **utilitarianism**.

Utilitarianism represents the most influential form of consequentialism. It can be analyzed as the combination of three core principles (Broome, 1991):

1. **Consequentialism:** Every moral choice, whether concerning acts, rules, or institutions, is ultimately justified by the **goodness of its consequences**.
2. **Sum-ranking:** The utility of any state of affairs is determined by the **sum total of all individual utilities** within it.
3. **Welfarism:** The goodness of a state of affairs depends solely on the **well-being** (or utility) of individuals affected by it.

In this framework, **happiness** or **well-being** is the only thing valuable **for its own sake**. All other goods (wealth, freedom, virtue) are valuable only insofar as they **promote happiness** (Mill, 1863). Utilitarianism thus asserts that actions are right in proportion as they tend to **promote happiness**, and wrong as they tend to produce the **reverse of happiness**.

Within utilitarianism, the **theory of value** — what counts as "good" — can be understood in **subjective** or **objective** terms.

(a) Subjective Theories

Classical hedonists such as **Jeremy Bentham**, **John Stuart Mill**, and **Henry Sidgwick** identify utility with **pleasure and the absence of pain**. For them, moral value derives from subjective experiences of pleasure and suffering.

Later economists, notably **Paul Samuelson**, reformulated this in behavioral terms through **revealed preference theory**, where welfare is inferred from individual choices — still a subjective, preference-based conception of the good.

(b) Objective Theories

By contrast, **Aristotle**, and later **Amartya Sen** and **Martha Nussbaum**, proposed **objective theories** of well-being, emphasizing the realization of human **capabilities** and **functionings**. The "good life" consists not merely in pleasurable experiences but in **developing and exercising one's human capacities**.

Thus, within consequentialist reasoning, there is ongoing debate about what the good to be maximized actually consists in: **subjective satisfaction** or **objective flourishing**.

Happiness, Well-Being, and Eudaimonia

In classical utilitarianism, **happiness** denotes the net balance of positive over negative feelings. The **net happiness** of an individual over any period can be described as their positive affect (pleasure, joy, satisfaction) minus their negative affect (pain, anxiety, discomfort), with each weighted by **intensity and duration**.

While "happiness" remains the most direct and widely understood term, broader notions such as **well-being** or **eudaimonia** capture richer dimensions of human flourishing. According to **Aristotle**, *eudaimonia* is not a passing emotional state but a **lifelong activity of living well**, guided by reason and virtue. Modern eudaimonic theories similarly hold that happiness involves **meaning, purpose, and personal growth**, not mere pleasure.

This broader interpretation aligns with **capabilities-based** accounts of welfare and complements the utilitarian emphasis on happiness by recognizing that **subjective satisfaction and objective fulfillment** are intertwined dimensions of human good.

From a social and policy perspective, utilitarianism maintains that happiness remains a **powerful guide for moral and political decision-making**. Collective well-being — understood as the promotion of happiness and the reduction of suffering — provides a **universal and measurable** standard for evaluating institutions and actions. As argued by **Yew-Kwang Ng** (1997) and **John Broome** (1991), happiness can be conceptualized, compared, and aggregated across individuals, thereby enabling rational moral and policy calculus. Even if imperfect, this framework offers a **pragmatic and human-centered** foundation for ethics, linking moral reasoning to real improvements in people's lives.

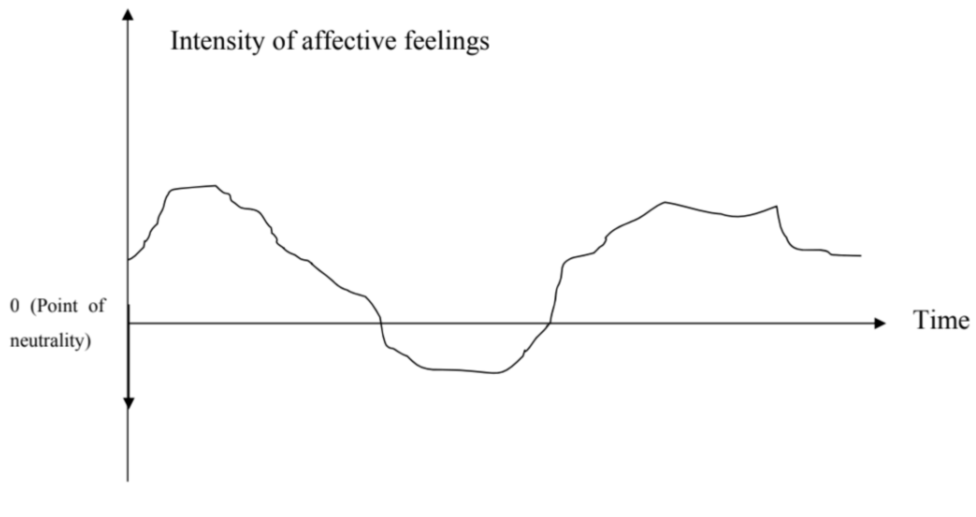


Fig. 1 Amount of happiness illustrated

Consequentialism provides a **systematic and outcome¢¢-oriented** approach to moral evaluation. It identifies the **good** as the foundational moral value, defines the **right** in terms of maximizing that good, and evaluates moral worth by reference to **results rather than intentions or rules**.

However, its completeness depends on the adequacy of its **theory of value** and **aggregation**. Utilitarianism offers one of the most enduring responses to this challenge by grounding moral evaluation in the **aggregate happiness and well-being of persons**. From Bentham's hedonism to Sen's capabilities, it represents an evolving effort to reconcile **individual experience** with **objective human flourishing**.

Ultimately, teleological ethics challenges us to look beyond duty to outcome, beyond rule to result, and beyond isolated acts to the broader **states of affairs** that define a better world. It endures as one of the most compelling traditions in moral philosophy precisely because it demands an answer to the fundamental moral question:

Does this make the world better for those who live in it?

2.3 From Utilitarianism to Welfare Economics: Aggregation, Efficiency, and the Limits of Collective Choice

Utilitarianism has been one of the most influential ethical doctrines in the formation of modern economics. Emerging from the moral philosophy of **Jeremy Bentham** and **John Stuart Mill**, it proposed that society should be organized to maximize the sum of individual happiness or utility. The utilitarian principle of maximizing aggregate welfare provided the moral and analytical foundation for classical and neoclassical economics, turning ethical reasoning into a **quantitative calculus** of social welfare.

Economists such as **James and John Stuart Mill**, and later **William Stanley Jevons**, **Alfred Marshall**, **Francis Edgeworth**, and **Arthur Pigou**, all shared the conviction that individual utilities could be **measured (cardinally)** and **compared across persons (interpersonally)**. This made it possible to think of society as a **sum of individual well-beings**, allowing economists to evaluate alternative policies by calculating the **net change in total utility**. If a policy A produced greater total happiness than policy B, utilitarianism implied that A was the morally and economically superior choice.

By the 1930s, the utilitarian foundations of welfare economics began to erode. Economists increasingly questioned the assumption that individual utilities were **cardinally measurable** or **interpersonally comparable**. The rise of **ordinal utility theory** — the idea that preferences can be ranked but not measured in absolute units — led to the rejection of the utilitarian aggregation framework.

In its place, economists adopted **Pareto efficiency** (or Pareto optimality) as a new, ostensibly value-free criterion of welfare. A state of allocation of resources is **Pareto optimal** if it is **impossible to make any individual better off without making someone else worse off**. A **Pareto improvement** occurs when at least one individual becomes better off while no one becomes worse off.

This redefinition of welfare sought to avoid ethical assumptions about comparing happiness across persons. Yet it came at a price: since most real-world policy changes **benefit some and harm others**, the Pareto criterion is **silent** in almost all interesting cases. Without a way to aggregate individual gains and losses, welfare economics lost its **normative force**.

The Two Fundamental Theorems of Welfare Economics

The Pareto framework was formalized in the **First and Second Fundamental Theorems of Welfare Economics**, which link market behavior to welfare outcomes.

1. **First Fundamental Theorem:** Under conditions of perfect competition, complete markets, and rational preferences, a competitive equilibrium will result in a **Pareto-efficient allocation** of resources.
   In other words, markets tend naturally toward efficiency.
2. **Second Fundamental Theorem:** Any Pareto-efficient allocation can be achieved by some **redistribution of initial endowments**, followed by competitive market exchange.
   This theorem provided a conceptual separation between **efficiency** (a technical problem) and **equity** (a distributive problem). Policymakers could, in theory, redistribute wealth and then rely on markets to achieve efficiency.

These theorems reflect the continuing utilitarian intuition that economics should aim for **maximum welfare**, but they operationalize it through **efficiency criteria** rather than direct utility maximization. The moral content of welfare economics thus became more implicit, even as it retained its utilitarian roots.

The limitations of the Pareto approach motivated attempts to reintroduce **utilitarian aggregation** without assuming cardinal utilities. The **Hicks–Kaldor compensation criterion** proposed that a policy change is **socially desirable** if the winners could, in principle, **compensate the losers** and still be better off — even if compensation does not actually occur.

This "potential Pareto improvement" restored a **decision-making rule** for cases involving trade-offs, while maintaining the formal apparatus of Pareto efficiency. However, it remained **implicitly utilitarian**: it assumes that increases in the utilities of the winners can morally outweigh decreases in those of the losers, even without explicit interpersonal comparisons. As such, the Hicks–Kaldor test reintroduced the **aggregation problem** it sought to avoid.

The Challenge of Collective Rationality: Arrow's Impossibility Theorem

In the mid-twentieth century, **Kenneth Arrow** (1951) exposed the deep logical limitations of any attempt to aggregate individual preferences into a coherent social choice function. His **Impossibility Theorem** (also known as the **General Possibility Theorem**) demonstrated that no social decision rule can simultaneously satisfy a set of apparently reasonable conditions — including **unrestricted domain**, **Pareto**

efficiency, **independence of irrelevant alternatives**, and **non-dictatorship** — when there are three or more options (**see BOX # 2.1**).

Arrow's result revealed the inherent **incoherence of utilitarian aggregation** once ordinal preferences replace cardinal utilities. The idea of a "social welfare function" summing individual utilities is incompatible with minimal conditions of collective rationality. In effect, Arrow's theorem marks the **mathematical end** of classical utilitarian welfare economics.

Sen's Liberal Paradox: The Limits of Paretian Liberalism

Building on Arrow's insight, **Amartya Sen** (1970) introduced the **Liberal Paradox**, also known as the **Impossibility of a Paretian Liberal**. Sen showed that even minimal respect for **individual freedom** — allowing each person to decide at least one matter of personal concern — is incompatible with the **Pareto principle** under certain conditions.

The paradox demonstrates that a society cannot simultaneously be:

1. **Pareto efficient**, and
2. **Perfectly liberal**, in granting individuals control over their personal choices.

Sen's result reveals a profound moral tension in welfare economics: **individual rights and social efficiency cannot always coexist**. This insight reopens the question that utilitarianism had sought to close — how to reconcile **individual autonomy** with **collective welfare**.

Beyond Utilitarianism: From Welfare to Capabilities

Sen's later work on the **capabilities approach** represents a partial return to the moral roots of utilitarianism, while rejecting its reductionism. Like Mill, Sen regards human well-being as the ultimate aim of social arrangements, but he defines it not in terms of subjective utility but in terms of **capabilities** — the real freedoms people have to lead lives they value.

This reorientation reframes welfare economics as a study of **human development** rather than market efficiency. It recovers the **normative concern** of classical utilitarianism — promoting human flourishing — while acknowledging the **plural and incommensurable** dimensions of well-being that resist aggregation.

The evolution of welfare economics reflects a continuing dialogue with utilitarianism. Classical economists sought to apply Bentham's moral arithmetic to the economy; neoclassical economists refined it into a formal theory of utility; modern theorists like Arrow and Sen revealed its internal contradictions. Yet the utilitarian aspiration — to organize society for the greatest good of its members — continues to animate economic thought.

From **Pareto efficiency** to **social choice theory**, the central challenge remains the same: how to balance **individual welfare**, **collective rationality**, and **moral fairness** in a world where utilities cannot be perfectly measured or compared. The history of welfare economics thus stands as both a **continuation** and a **critique** of utilitarianism — a search for a moral language adequate to the complexity of human well-being.

BOX # 2.1: From Condorcet's Paradox to Arrow's Impossibility Theorem

One of the most profound challenges in collective decision-making lies in the tension between **individual preferences**and **collective rationality**. This tension was first clearly exposed in the 18th century by the Marquis de **Condorcet**, and later formalized by **Kenneth Arrow** in his 1951 work *Social Choice and Individual Values*. Both thinkers grapple with a fundamental question:

How can we derive a consistent and fair social choice from the diverse preferences of individuals in society?

Condorcet's Voting Paradox

Condorcet observed that even when individuals have clear and consistent preferences, the process of **majority voting**can produce **collectively inconsistent** outcomes.

Imagine a simple example with **three voters (1, 2, 3)** and **three alternatives (x, y, z)**, representing possible social states or policies.

Voter	1st Choice	2nd Choice	3rd Choice
1	x	y	z
2	y	z	x
3	z	x	y

Now, let's use **pairwise majority voting**:

- Between **x** and **y**: Voters 1 and 3 prefer x → **x beats y**
- Between **y** and **z**: Voters 1 and 2 prefer y → **y beats z**
- Between **z** and **x**: Voters 2 and 3 prefer z → **z beats x**

The collective ranking is therefore **cyclical**:

x is preferred to y,
y is preferred to z,
but z is preferred to x.

This violates **transitivity**, the basic requirement of rational preference ordering. Even though every individual's ranking is perfectly rational, the collective outcome is not. This is known as the **Condorcet paradox** — a situation in which **majority rule produces inconsistency** and no clear social winner exists.

From Condorcet to Arrow

Condorcet's paradox revealed a structural flaw in majority decision-making, but Kenneth **Arrow** showed that the problem is not confined to majority rule — it is **universal**. In his *Impossibility Theorem*, Arrow generalized the paradox to **all conceivable decision rules** that aim to aggregate individual preferences into a single, coherent **social welfare function**.

Arrow's starting point was the idea that society must choose among a set of possible states — *x,y,z,…*x,y,z,…. Each individual *i*i has a preference ranking *Ri*Ri over these states (for example, x > y > z). The goal is to define a **social ranking** *R*R — an ordering of these states that represents society's overall preference — based on all individual rankings.

Arrow's Axioms

Arrow proposed that any reasonable social choice rule should satisfy a few apparently modest and democratic conditions:

1. **Unrestricted Domain (U):** The rule must work for *any* logically possible set of individual preferences.
2. **Pareto Principle (P):** If *everyone* prefers x to y, then society should prefer x to y.
3. **Independence of Irrelevant Alternatives (I):** The social ranking between x and y should depend *only* on how individuals rank x and y — not on other irrelevant options.
4. **Non-Dictatorship (D):** There should be *no single individual* whose preferences always determine the social preference, regardless of others' views.

Each of these axioms seems reasonable. Together, they define what we might consider a **fair and rational democracy**: it respects everyone's preferences, treats options independently, and avoids dictatorship.

The Impossibility Theorem

Arrow's **General Possibility Theorem** states that:

If there are at least three social states and a finite number of individuals,
no social welfare function can satisfy all four axioms (U, P, I, and D) at once.

In other words, **a perfectly fair and rational collective decision rule is impossible**. Any rule that respects unrestricted preferences, the Pareto condition, and independence will end up being **dictatorial** — it must give decisive power to one individual.

This conclusion is deeply unsettling. It implies that there is an inherent tension between **democratic fairness** and **logical consistency** in collective decision-making. As Arrow himself put it, "We are forced to the conclusion that there is no means of forming a social choice without either sacrificing rationality or imposing dictatorship."

Arrow's theorem generalizes the Condorcet paradox: while Condorcet showed that **majority rule** can fail, Arrow demonstrated that **every rule** is subject to similar inconsistencies, unless we relax one of the axioms. The theorem thus exposes a profound **ethical and epistemic limit** in the utilitarian dream of aggregating individual welfare into a coherent social welfare.

Economically, this undermines the idea of a purely objective or mechanical **"social welfare function"** that could guide policy. Politically, it forces us to acknowledge that **collective choice is unavoidably plural and conflictual** — that social rationality cannot be reduced to individual preferences.

Arrow's work inspired generations of thinkers to seek ways around the impossibility result:

- **Amartya Sen** showed that allowing for **interpersonal comparisons of welfare** and introducing **rights or freedoms** could restore some coherence to social choice, though at the cost of pure Pareto efficiency.

- **John Rawls** rejected preference aggregation altogether, grounding justice instead in principles chosen under fair conditions.
- The **Capabilities Approach**, again developed by Sen (and Nussbaum), redefines social evaluation not in terms of preferences but in terms of **what individuals are actually able to do and be**.

Thus, Arrow's impossibility theorem does not close the debate — it **transforms** it. It forces economists and philosophers to rethink what it means to aggregate welfare and to justify collective choices in pluralist societies. From Condorcet's cyclical majorities to Arrow's general impossibility, the lesson is the same: **collective rationality cannot simply be derived from individual rationality**. The path from personal preferences to social decisions is not logically straightforward, but normatively complex. Arrow's theorem remains one of the most elegant and humbling reminders that **democracy, fairness, and consistency cannot always coexist** — and that any viable theory of justice must navigate the trade-offs among them.

BOX # 2.2: Sen's Liberal Paradox — The Impossibility of a Paretian Liberal

Kenneth Arrow had shown that it is **impossible to aggregate individual preferences** into a coherent and fair social ranking that satisfies a few intuitive democratic axioms.
Amartya **Sen** — working within the same tradition of **social choice theory** — pushed this logic further. He asked:

What happens when we introduce *individual rights* or *liberties* into this system?
Can a society respect both **freedom** and **efficiency** (the Pareto principle) at the same time?

The answer, surprisingly, is **no**.

Sen's *"Liberal Paradox"* demonstrates that even a minimal requirement of **individual freedom** — the right of each person to decide in one small sphere of personal choice — can conflict with the **Pareto criterion**, which requires that social choices reflect unanimous improvements in welfare.

Thus, just as Arrow showed the **impossibility of a purely fair and consistent democracy**, Sen revealed the **impossibility of a society that is both perfectly liberal and perfectly efficient**.

Paradox

To see how the paradox works, Sen defines two principles:

1. **Minimal Liberalism:** Every individual should have at least some sphere of personal freedom — decisions that concern only themselves and should not be interfered with by others.
2. **Pareto Efficiency:** No social arrangement should be rejected if it makes at least one person better off without making anyone else worse off.

Sen shows that these two principles — liberty and efficiency — can be logically incompatible, even in the simplest possible society of two individuals and one decision.

Consider two individuals, **Prude** and **Lewd**, and the choice of who, if anyone, should read *Lady Chatterley's Lover*.

- **Prude** finds the book offensive and would prefer *not* to read it. However, he is even more disturbed by the thought of **Lewd** reading and enjoying it.
- **Lewd**, on the other hand, loves the book, but derives even more pleasure from imagining the prudish Prude being forced to read it.

Their preferences can be summarized as follows:

Situation	Prude's Ranking	Lewd's Ranking
(1) Prude reads, Lewd does not	**Best**	**Best**
(2) Lewd reads, Prude does not	**Worse**	**Worse**

From a **liberal** standpoint, the natural arrangement is for **Prude not to read** (since he dislikes the book) and **Lewd to read** (since he enjoys it). Each person controls their own private decision — a perfect embodiment of **minimal liberalism**.

However, from a **Pareto** standpoint, this outcome is inefficient. Both Prude and Lewd would prefer the situation in which **Prude reads** and **Lewd does not**. For society to reach the Pareto-optimal outcome, Prude must be coerced into reading and Lewd forbidden to do so.

Criterion	Outcome	Description
Minimal Liberalism	Prude (not read), Lewd (read)	Respects freedom, but inefficient
Pareto Efficiency	Prude (read), Lewd (not read)	Efficient, but violates freedom

The paradox shows that the liberal commitment to individual autonomy and the utilitarian commitment to social welfare maximization cannot both be satisfied. Respecting personal liberty may lead to inefficiency; achieving efficiency may require coercion.

Sen's paradox parallels a well-known result in welfare economics: **markets fail to achieve Pareto efficiency in the presence of externalities.**
In a perfectly competitive market, efficiency is reached only if each person's preferences concern *their own*consumption. But when one person's preferences depend on another's actions — as with pollution, overfishing, or in this case, moral disapproval — the smithian invisible hand fails.

In Sen's example, **Prude's and Lewd's preferences are interdependent**. Each cares about what the other does. This interdependence functions as a **moral externality**, breaking the independence of private spheres assumed by both liberal and utilitarian theories. It becomes impossible to ensure both perfect freedom and Pareto efficiency.

Interpretation

Sen's paradox reveals a deeper truth about **collective ethics and liberal democracy**. Whereas **Arrow's Impossibility Theorem** showed that *democracy and rationality* can conflict, **Sen** demonstrated that *freedom and efficiency* can also be irreconcilable under purely preference-based reasoning.

In both cases, the root of the problem lies in **treating social choice as a mechanical aggregation of individual preferences.** Preferences are not value-neutral: they may express moral judgments, interpersonal concerns, or envy and resentment. Once such preferences are admitted, the logic of welfare economics collapses — respecting individual rights may come at the cost of efficiency.

Thus, Sen's *Impossibility of a Paretian Liberal* shows that no social decision procedure can perfectly reconcile **individual freedom** and **collective welfare** if we assume that preferences fully capture what matters ethically. The paradox exposes the **limits of preference-based utilitarianism**: even perfect information and rationality cannot ensure justice or harmony between liberty and welfare.

Sen's insight invites a profound philosophical shift. Like Arrow before him, he reveals the incompleteness of welfare economics and utilitarian ethics when they rely solely on preference aggregation. But where Arrow diagnosed the conflict between *collective rationality* and *democracy*, Sen uncovers the conflict between *liberty* and *efficiency*.

In both cases, social choice cannot be reduced to counting or comparing preferences. Justice requires public reasoning about rights, fairness, and the moral significance of what individuals value.

Sen's paradox thus stands as a foundational argument for broadening ethics beyond utility — toward an understanding of justice grounded in **freedom as capability**, **reasoned values**, and **respect for persons as moral agents**, not merely preference-bearers.

2.4 From Pleasure to Dignity: Bentham, Mill, and the Evolution of Utilitarianism

Modern utilitarianism, one of the most influential moral and political theories of the nineteenth century, was born from the Enlightenment faith in reason, progress, and reform. Its founding premise is deceptively simple: human action, institutions, and laws should aim at maximizing happiness and minimizing suffering. Yet this simplicity conceals a radical redefinition of ethics itself — away from divine command or inherited custom and toward a **secular, rational calculus of welfare** (Bentham, 1789/1996).

Jeremy Bentham and **John Stuart Mill** stand as the two towering figures of this tradition. Bentham formulated its logic in the language of **hedonistic empiricism**, envisioning a moral science grounded in measurable pleasure and pain. Mill, his intellectual heir, sought to deepen and humanize the doctrine, defending liberty, individuality, and the cultivation of higher faculties as integral to true happiness. Between them unfolds one of the most significant transformations in modern moral philosophy: the movement from **quantitative to qualitative utilitarianism**, from pleasure as arithmetic to happiness as human flourishing (Mill, 1861/2001).

Jeremy Bentham

Born in 1748, **Jeremy Bentham** was a precocious child of the Enlightenment. Educated at Oxford and trained as a lawyer, he devoted his life to reforming the English legal system and society at large. A prolific writer and tireless advocate of reason, transparency, and equality, Bentham corresponded with figures such as **Adam Smith** and inspired a generation of reformers, including **James Mill**, **John Stuart Mill**, **John Austin**, and **Robert Owen** (Schneewind, 1990).

Bentham's ethical philosophy begins from a simple observation about human nature:

"Nature has placed mankind under the governance of two sovereign masters, pain and pleasure" (Bentham, 1789/1996, p. 1).

Everything we do, he argued, is motivated by the desire to seek pleasure and avoid pain. These experiences are not only psychological facts but also **moral standards**. To determine what is right or wrong, we must ask whether an action tends to increase or diminish the sum total of happiness. This is the **principle of utility**.

Bentham proposed a "felicific calculus" by which pleasures and pains could, at least in principle, be measured and compared according to their **intensity, duration, certainty, and extent**. On this view, all sources of pleasure — bodily or mental, refined or vulgar — are fundamentally of the same kind. The Latin maxim *de gustibus non disputandum*captures Bentham's attitude: there is no disputing about tastes (Bentham, 1789/1996).

This **moral egalitarianism** was both Bentham's strength and his limitation. By treating all pleasures as commensurable, he offered a framework of justice blind to status, tradition, or prejudice. It was this radical impartiality that inspired his calls for **women's equality**, **animal rights**, and even the **liberalization of homosexual acts** (Hart, 1970). Yet it risked **moral flattening**: if all pleasures are measured by the same scale, there seems to be no ethical difference between the joy of music and the thrill of cruelty.

In his personal life, Bentham embodied his philosophy's confidence in reason and reform. He left instructions that his body be dissected and preserved as an **"auto-icon"**, which still sits today at **University College London** — a literal symbol of Enlightenment rationalism's self-assurance.

John Stuart Mill: Liberty and the Education of Humanity

John Stuart Mill (1806–1873) inherited Bentham's philosophy and the rigorous education of his father, **James Mill**, a close collaborator of Bentham. Trained from infancy in Greek, Latin, logic, and political economy, the young Mill was, in his own words, a "manufactured philosopher" (Mill, 1873/1998). His early intellectual formation led to a profound **mental crisis** in adolescence: the discovery that reason and logic alone could not give life meaning. This crisis, and his recovery through literature and friendship, shaped his later philosophy, emphasizing **emotional and moral depth** along side rational calculation.

Mill's professional life combined public service and scholarship. For twenty-five years he served as a **colonial administrator at the East India Company**, and in later years became a **Member of Parliament** for the Liberal Party. As a reformer, he championed

women's rights, labor unions, cooperative ownership, and the **extension of suffrage**(Cranston, 1967).

In *On Liberty* (1859/1978), Mill articulates a powerful defense of individual freedom:

"Over himself, over his own body and mind, the individual is sovereign" (p. 13).

People should be free to act as they wish so long as they do not **harm others**. Yet this principle appears to conflict with utilitarianism's collective calculus: if the majority's happiness increases by suppressing a minority's lifestyle or religion, why not act? Mill answers through **long-term utility**: a society that silences dissent stifles innovation, moral growth, and intellectual progress. Liberty is not an end in itself but a **means to enduring social happiness**.

Mill's Departure: Qualitative Pleasures and Human Dignity

Mill further revises Bentham's doctrine in *Utilitarianism* (1861/2001) by distinguishing **higher and lower pleasures**. While Bentham treated all pleasures as commensurable, Mill argues that intellectual, moral, and aesthetic pleasures are intrinsically superior to mere sensual gratification:

"It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied" (p. 18).

Higher pleasures are not superior because people prefer them, but because they engage the **higher faculties** of human nature. This introduces an **ethical standard independent of preference satisfaction**, partially transcending classical utilitarianism. Happiness, for Mill, is not merely a state of pleasure but a **mode of existence** — the flourishing of a free, rational, and morally cultivated individual.

This shift also informs Mill's political thought. His advocacy for **women's rights**, cooperative labor, and universal suffrage reflects a utilitarianism infused with **justice and empathy**. Liberty and moral development become central to happiness, emphasizing that the **character of persons matters** in evaluating the good of society.

Reconciling Utilitarianism and Individual Rights

A key tension in Mill is reconciling **utilitarian calculus with individual rights**. Bentham's framework could, in principle, justify sacrificing the interests of a minority if it increased overall happiness. Mill, however, insists that respecting individual freedom often promotes long-term utility. By allowing dissenting opinions and lifestyles to exist, society ensures innovation, moral growth, and the continual testing of ideas. In this sense, Mill transforms Bentham's quantitative utilitarianism into a **humanistic ethics**, where liberty, character, and moral development are essential components of the good life (Schneewind, 1990).

Two Faces of Utilitarian Modernity

Bentham and Mill represent complementary faces of utilitarian modernity. **Bentham's utilitarianism** is rational, egalitarian, and mechanistic, seeking to maximize welfare through impartial calculation. **Mill's utilitarianism** is humanistic, qualitative, and aspirational, harmonizing happiness with freedom, dignity, and self-realization. Bentham's world was one of **social engineering**; Mill's was one of **moral education**.

Where Bentham sought a scientific measure of pleasure, Mill sought the cultivation of character. Yet both shared a radical commitment: the good society is not founded on divine command or aristocratic privilege, but on the **well-being of all sentient beings**, measured through the expansion of human capacities and the minimization of suffering. From Bentham's arithmetic to Mill's philosophy of dignity, utilitarianism evolves from the **science of pleasure** to a **moral project of human flourishing**.

2.5 Justice and Utility in Mill's Utilitarianism: Reconciling Moral Rights with the Greatest Happiness Principle

From its inception, utilitarianism faced a profound moral challenge: **how can a theory that judges rightness solely by consequences account for the binding, impartial demands of justice?** If morality consists only in maximizing overall happiness, why should individuals respect rights or fairness when violating them might yield greater utility?

This question becomes acute in **John Stuart Mill's *Utilitarianism*** (1861/2001), particularly in **Chapter V, "On Justice"**, where Mill confronts the charge that utilitarianism cannot explain the moral force of justice. Justice seems to involve **absolute duties and legitimate claims** that resist the fluid calculations of utility. Yet, for Mill, this appearance conceals a deeper harmony: justice, rightly understood, is not opposed to utility but rather its most vital and indispensable expression.

Mill's project in Chapter V is both **conceptual and moral**: to clarify what we mean by *justice*, to trace its psychological and social origins, and to demonstrate that its authority ultimately derives from the **general utility**. His analysis also anticipates a distinction central to later utilitarian thought — between **act-utilitarianism** (judging each action by its consequences) and **rule-utilitarianism** (judging rules or institutions by their tendency to maximize happiness). Justice, Mill suggests, belongs to the latter domain.

The Problem of Justice: Its Distinctive Moral Character

Mill begins Chapter V by noting that justice appears "to be the name for certain classes of moral rules which concern the essentials of human well-being more nearly, and are therefore of more absolute obligation, than any other rules for the guidance of life" (*Utilitarianism*, V.1). Justice seems to possess two unique features:

1. **Stringency** – its demands are felt as absolute and unconditional; and
2. **Correlativity** – it involves the rights of others, giving rise to legitimate claims that must be respected.

For instance, the duty not to punish the innocent or to distribute goods fairly seems morally binding irrespective of consequences. This intuition appears to contradict utilitarianism's **teleological flexibility**: if consequences alone determine rightness, why not violate justice when doing so increases aggregate happiness? This tension underlies classic objections to utilitarianism as a "cold" and "expedient" doctrine that cannot secure moral rights (Rawls, 1971).

Mill accepts the emotional reality of this intuition. The sense of justice, he argues, involves a **powerful sentiment of resentment** toward harm or injustice done to oneself or others. Yet this sentiment is not an innate moral faculty; rather, it is a **socially cultivated feeling** rooted in our instinct for self-preservation and our capacity for sympathy. Justice, then, expresses an evolved form of social utility — a sentiment whose general observance promotes collective security and trust.

Mill's Psychological and Social Analysis of Justice

Mill's first move in reconciling justice with utility is to **demystify the moral feeling of justice**. He distinguishes its **subjective origin** from its **objective justification**.

- **Subjectively**, the sense of justice derives from two natural impulses:
  (a) the desire for retaliation or vengeance against harm, and
  (b) the sympathetic identification with others that extends this feeling beyond the self.

Over time, moral education and social experience transform these impulses into a stable sentiment supporting impartial fairness.

- **Objectively**, what gives this sentiment moral authority is not its strength but its **social usefulness**: the general observance of justice is indispensable for human well-being. "Justice," Mill writes, "is a name for certain moral requirements, which, regarded collectively, stand higher in the scale of social utility, and are therefore of more paramount obligation, than any others" (V.32).

Thus, justice is not an independent moral domain but a **subset of utility** — those rules whose violation would most gravely undermine social happiness. The *appearance* of moral absoluteness arises because justice concerns **security**, the most basic condition for collective life. A society where people cannot trust others to respect their rights cannot sustain any stable form of happiness.

Mill's key reconciliation comes with a famous statement:

"Justice is a name for certain moral requirements which, regarded collectively, stand higher in the scale of social utility and are therefore of more paramount obligation than any others." (V.32)

Justice, in this sense, is not external to utilitarianism but **its most sacred part**. The rules of justice — prohibitions on harming others, requirements of fairness, protection of rights — are those whose general observance yields **maximal long-term utility** by securing the basic trust and reciprocity upon which happiness depends.

Violating justice may produce short-term gain but undermines the moral fabric that sustains social cooperation. Mill's reasoning here is proto-rule-utilitarian: rather than asking whether a *particular act* maximizes utility, we must ask whether a **rule or principle** permitting such acts would increase happiness if generally followed.

For example, punishing the innocent might in one case appease public anger, but a rule allowing such punishment would destroy public confidence in justice, yielding immense suffering. Therefore, the utilitarian defense of justice appeals to **general consequences** — to the indispensable social role of rules protecting individual security.

Mill's treatment of justice anticipates the later distinction between **act-** and **rule-utilitarianism**, though he never uses these terms explicitly.

- **Act-utilitarianism** judges each action directly by its consequences for overall happiness.
- **Rule-utilitarianism** evaluates rules or practices by their general tendency to promote happiness and holds that individuals should follow such rules even when a particular exception might appear beneficial.

In Chapter V, Mill presents a rule-utilitarian framework: the principles of justice are understood as moral rules whose general observance is essential to collective security and thus to overall well-being. They must therefore be treated as binding in nearly all circumstances. However, they do not possess absolute, deontological authority: exceptions are permitted only in rare cases where violating the rule is the only way to avert a greater harm to the security and happiness of the community. Ultimately, utility remains the foundational standard that justifies both the rules and the rare departures from them. (see Box # : The Trolley Problem).

Thus, Mill's utilitarianism incorporates a **layered moral hierarchy**:

1. The *ultimate principle* remains utility — the promotion of the greatest happiness.
2. But *secondary moral rules* (e.g., keep promises, do not harm, respect rights) derive their authority from their stable utility across time and society.
3. Justice occupies the highest level among these secondary principles because it secures the **preconditions for happiness** — trust, stability, and mutual protection.

This hierarchy transforms utilitarianism from a mechanical calculus of outcomes into a **structured moral framework** grounded in moral psychology, institutional trust, and long-term social utility.

Justice, Rights, and the Limits of Expediency

Mill's account also illuminates the moral status of **rights**. He defines a right as a valid claim society ought to protect by force of law or opinion. The concept of rights thus becomes a **social guarantee of utility**: individuals have rights to those things that it is for the general good that they should possess securely.

This formulation bridges utilitarianism and liberalism. Rights are not natural or pre-political, but neither are they mere conventions. They are **institutions justified by their contribution to the greatest happiness**, especially through securing the conditions of mutual respect and freedom.

In this sense, justice imposes limits on expedient utilitarian reasoning. Because the rules of justice protect the essential interests of individuals, they must not be overridden for minor gains. Only when violating a right prevents a catastrophe that would destroy even greater security can exceptions be justified. Thus, Mill's utilitarianism is **teleological but constrained**, a moral system governed by principles of utility mediated through stable rules and moral sentiments.

Mill's reconciliation of justice and utility marks a **maturation of utilitarian ethics**. Bentham's system had been criticized for its "felicific calculus," which seemed to reduce moral reasoning to arithmetic and to ignore the qualitative texture of moral life. Mill restores to utilitarianism a sense of moral depth and institutional realism.

Justice, on Mill's view, emerges from the interplay of **human psychology, social cooperation, and long-term welfare**. The sense of moral obligation is neither an illusion nor an independent categorical imperative; it is the emotional correlate of our rational recognition that some moral rules are foundational for the happiness of all.

This view anticipates later developments in **rule consequentialism** and **contractual utilitarianism**, as seen in the works of R. M. Hare, J. J. C. Smart, and Brad Hooker. It also aligns with contemporary theories that interpret rights as **rules of justified constraint** within a consequentialist framework.

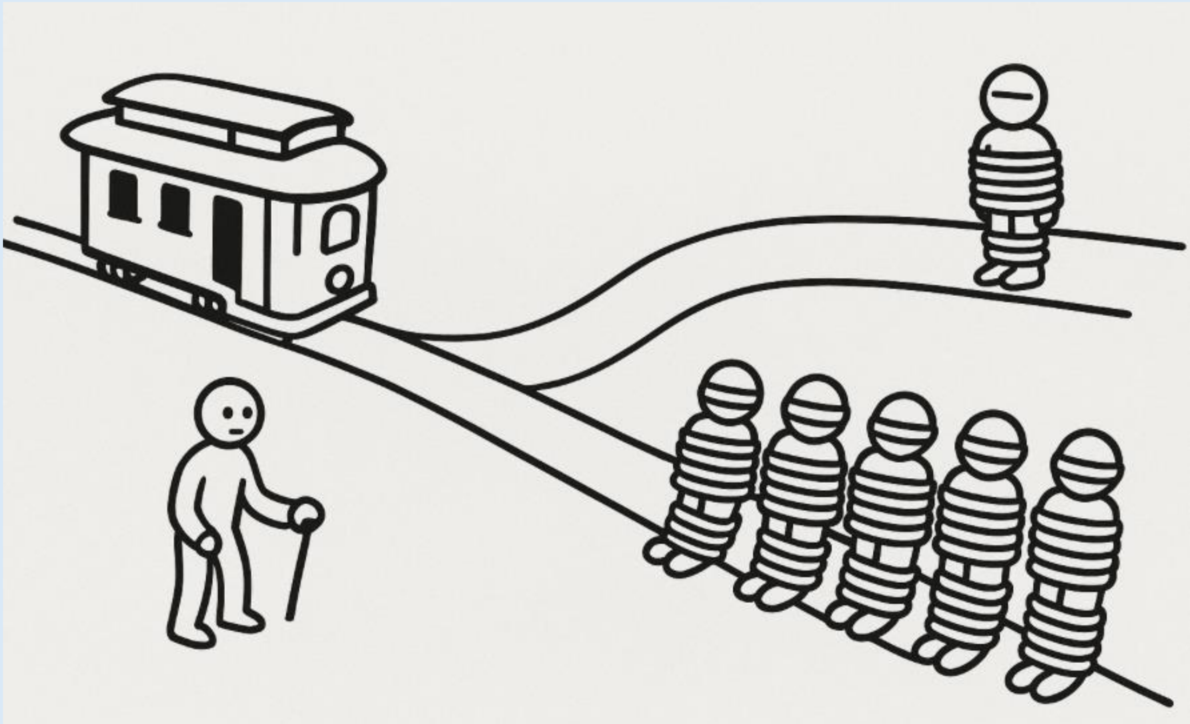
Mill's reconciliation of justice and utilitarianism in Chapter V of *Utilitarianism* remains one of the most sophisticated attempts to integrate **moral rights with consequentialist ethics**. Far from being incompatible, justice and utility, for Mill, are conceptually intertwined:

- Justice names those moral rules that protect the most vital conditions of human happiness.
- The sense of justice, though rooted in emotion, finds its moral justification in general utility.
- Rights are justified expectations society must protect because their observance maximizes long-term welfare.

In this synthesis, Mill transforms utilitarianism from a calculus of pleasures into a **comprehensive moral philosophy**, one that unites the value of happiness with the dignity of persons. Justice, rather than limiting utilitarianism, becomes its **moral culmination** — the principle that secures the very possibility of collective happiness through the protection of individual worth.

BOX # 2.3: The Trolley Problem

In its standard form, the trolley problem goes like this: a runaway trolley is heading down a track where it will kill five people tied up and unable to move. The driver is unconscious, and no one can stop the trolley directly. You are standing next to a lever that can divert the trolley onto another track — but on that track there is one person tied up.



Should you pull the lever, sacrificing one to save five?

The case sets two moral intuitions against each other:

- **The utilitarian intuition**: Pull the lever, because saving five lives maximizes overall welfare.
- **The deontological intuition**: Do not pull the lever, because intentionally killing an innocent person violates a fundamental right, regardless of consequences.

On a utilitarian reading of Mill:

- **Act utilitarianism**: Evaluate the act by its immediate consequences.
  → Pulling the lever seems morally required (five saved vs. one lost).
- **Rule-guided utilitarianism**: Obey rules that, in general, best promote security and happiness.
  → Because deliberately harming innocents typically erodes trust and social stability, one might not pull the lever.

However, the trolley case presents **further nuance** that complicates the idea that pulling the lever simply "violates a right." The danger exists independent of the agent, and someone will die no matter what. The agent is not introducing a new harm but attempting to **minimize unavoidable losses** caused by circumstances beyond anyone's responsibility. The single victim's right to security has already been compromised by the runaway trolley, not by the bystander's intervention. From this perspective, the agent may not be violating rights but **redistributing risk** under tragic constraints.

This is exactly the kind of tragic scenario Mill aims to clarify in Chapter V. Justice represents rules essential to collective security and thus to the general happiness. These rules are normally binding because they protect vital expectations: that people will not be arbitrarily harmed. Yet Mill also insists that **utility remains the ultimate standard** — in catastrophic exceptions where rights have already broken down, the moral task may be to salvage as much security as possible.

So the trolley problem does not simply show justice and utility in conflict; it reveals their **conditional relationship**. Mill's position suggests:

- **In normal life**: Rights must be upheld strictly to preserve happiness-enabling security.
- **In extraordinary emergencies**: Utility may justify exceptions precisely to preserve the greatest amount of security remaining.

The thought experiment thus highlights the enduring challenge in Mill's project: balancing **respect for each person's security** with **the imperative to reduce overall harm** when not all rights can be protected. It shows why Mill believes utility is the *foundation* of justice — not because justice can be ignored, but because its value lies in how it contributes to the flourishing of human life together.

2.6 Mill's Justification of Utilitarianism: Reconciling Justice, Liberty and the Greatest Happiness

John Stuart Mill stands at a crucial crossroads in the history of moral philosophy. Inheriting Jeremy Bentham's doctrine of utility — the idea that right action is that which produces "the greatest happiness of the greatest number" — Mill sought to defend and humanize it. His project was twofold: to show that utilitarianism could provide a rational foundation for morality, and to demonstrate that it could accommodate the moral sentiments of justice, rights, and individuality that seemed to elude Bentham's more mechanical calculus of pleasure. In *Utilitarianism* (1861) and *On Liberty* (1859), Mill transformed the utilitarian tradition from a quantitative hedonism into a richer ethical system grounded in the development of human capacities, the cultivation of moral character, and the long-term welfare of society. This section examines how Mill justifies utilitarianism as a moral theory, how he reconciles it with the demands of justice, and how his synthesis anticipates later liberal theories such as Rawls's *justice as fairness* and Sen's *capability approach*.

The Challenge of Justification

In *Utilitarianism* (1861), John Stuart Mill set himself a daunting philosophical task: to defend the "greatest happiness principle" as the **foundation of morality**. According to this principle, *actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness*.

But how can this be justified? Why should individuals regard the promotion of happiness — especially the **general happiness** — as the supreme moral end? Moral philosophers before and after Mill have struggled with this question. Unlike a mathematical theorem, morality cannot be proven by deduction. It appeals instead to **experience, reflection, and human motivation**. Mill's "proof" of utilitarianism, presented in Chapter IV of *Utilitarianism*, remains one of the most debated arguments in moral philosophy.

The Nature of Moral Proof: From Facts of Desire to the Principle of Utility

Mill begins his justification by asking what kind of proof moral principles can have. He rejects any **a priori** or transcendental foundation. Moral principles, like first principles of reasoning, are **ultimate ends** — they cannot be demonstrated in the same way empirical facts can. Still, we can examine them empirically: "Questions of ultimate ends are not amenable to direct proof," he writes, "but the only proof that something is desirable is that people actually desire it."

This statement reveals Mill's **empirical and psychological foundation** for ethics. Just as the only evidence that something is visible is that people see it, the only evidence that something is desirable is that people desire it. Mill's reasoning is thus *analogical*, not deductive: it infers from human motivation to moral value.

He proceeds in two steps:

1. **Each person desires his or her own happiness.**
   → Therefore, happiness is a good to each person.
2. **The general happiness is nothing but the sum of individual happiness.**
   → Therefore, the general happiness is a good to the aggregate of all persons.

From this, Mill concludes that happiness is the *only* thing desirable as an end, and everything else (virtue, wealth, power) is desirable only as a means to or part of happiness.

The Alleged Fallacy: From "Is Desired" to "Ought to Be Desired"

Critics have often accused Mill of committing a **naturalistic fallacy** — inferring that because something *is desired*, it *ought to be desired*. G. E. Moore later called this the "naturalistic mistake," arguing that no factual statement about desire can entail a normative conclusion about what is good.

However, this objection misunderstands Mill's strategy. Mill does not claim that desire *logically entails* moral obligation. Rather, he argues that the *only kind of evidence available* for something being desirable (in the sense of "worthy of desire") is that it is, as a matter of fact, desired under conditions of full experience and reflection. In other words, Mill offers a **psychological justification** for the ultimate end of morality: if all moral reasoning must terminate in something we take as valuable for its own sake, the only plausible candidate — supported by universal human motivation — is happiness.

Thus, Mill's proof is not a deduction from "is" to "ought" but an **empirical grounding of value** in human nature. It is a form of moral naturalism: what humans stably and universally pursue under conditions of reasoned reflection provides evidence of what is good.

From Individual to General Happiness

The second step of Mill's argument is often misunderstood. Critics object: even if each person desires his or her own happiness, it does not follow that everyone desires the happiness of all. Why, then, should *I* care about *others'* happiness?

Mill's answer lies in **moral development and social psychology**. Human beings, he argues, are naturally social creatures endowed with sympathy and capable of identifying with the happiness of others. Through education, moral cultivation, and social experience, individuals come to internalize a sense of unity with their fellow beings — a feeling that one's own good is bound up with the good of others. This "social feeling of mankind" becomes, in Mill's view, "the ultimate sanction of the greatest happiness morality."

Thus, the transition from individual to collective happiness is not a logical inference but a **psychological and moral transformation**. As civilization advances, people increasingly realize that their happiness depends on cooperation, trust, and mutual respect. Acting to promote the general happiness becomes not a sacrifice of self-interest but an expression of a **higher form of self-interest** — a cultivated identification with humanity.

In this way, Mill integrates **egoism and altruism** within a unified moral psychology. The moral agent's motivation to promote general happiness rests on both natural sympathy and rational reflection on the conditions of personal flourishing in society.

Why People Ought to Maximize Social Utility

So, why *should* people act to maximize social utility? Mill's justification combines **empirical realism**, **psychological insight**, and **moral idealism**:

1. **Empirically**, happiness is the only thing people ultimately desire for its own sake. Everything else — wealth, fame, virtue — is valued because it contributes to happiness. Therefore, happiness is the only intrinsic good.
2. **Socially**, human beings are interdependent. The stable pursuit of one's own happiness requires living in a cooperative and secure society. Justice, trust, and moral rules are conditions of general happiness and, therefore, of one's own.
3. **Morally**, reason and moral education expand our sympathies beyond ourselves. The "social feeling of mankind" transforms the pursuit of happiness from a private goal into a universal one. We come to value others' happiness as part of our own moral nature.

Therefore, people *ought* to act to maximize social utility because doing so expresses the most enlightened and complete form of their own nature as rational and social beings. In Mill's vision, morality is not an external constraint but the rational harmonization of personal and collective happiness.

The Role of Virtue and the Internal Sanctions of Morality

Mill also addresses the motivational problem: even if we *accept* that we should maximize happiness, what makes people actually *want* to do so?

He distinguishes between **external sanctions** (social approval, punishment, divine reward) and **internal sanctions** (the feeling of conscience and moral sentiment). The latter, he argues, are the true foundations of moral obligation. Through moral education, individuals internalize the connection between their own happiness and the happiness of others, developing an **internal feeling of duty** toward utility. This feeling becomes a psychological force as strong as any natural impulse.

Moreover, Mill expands utilitarianism to include **virtue** as part of happiness. Once individuals learn to value virtue — honesty, benevolence, courage — for their own sake, their desire for virtue becomes one with their desire for happiness. This integration resolves the apparent conflict between self-interest and morality: virtue contributes to happiness not as a mere means but as a **constituent part** of it.

Thus, Mill offers a **progressive vision of morality**: utilitarianism is not a cold calculus but an ethical ideal guiding humanity toward a state where the happiness of each is bound up with the happiness of all. In this sense, acting to maximize social utility is not an alien demand but the realization of our best and most human capacities — reason, sympathy, and moral growth.

Mill justifies utilitarianism by appealing to the empirical structure of human motivation and the moral evolution of society. People should act to maximize social utility because:

- Happiness is the only ultimate end of action;
- The happiness of each depends on the happiness of all; and
- Acting for the general good fulfills our rational and moral nature.

For Mill, the principle of utility is not an arbitrary command but a **rational reflection of what we already, by nature, desire**. To act for the greatest happiness is to act in accordance with our highest faculties — as autonomous, sympathetic, and social beings.

Mill vs. Kant

The contrast between **John Stuart Mill** and **Immanuel Kant** offers an interesting philosophical insight into two radically different ways of grounding moral obligation. Both philosophers sought to provide a rational foundation for ethics, but they began from different assumptions about **human motivation**, **reason**, and **the nature of moral value**.

(a) The Source of Moral Authority

For **Kant**, the moral law derives from **pure reason**. The only thing unconditionally good is a *good will*, and the moral law commands categorically: “Act only according to that maxim whereby you can at the same time will that it should become a universal law.” Moral obligation is **a priori** and autonomous: it binds rational agents regardless of their inclinations, desires, or consequences. What makes an action right is not its outcome but its conformity to the moral law as dictated by reason.

For **Mill**, by contrast, the source of moral obligation is **empirical and psychological**. Morality arises from human nature — from the fact that people desire happiness and can come to care about others’ happiness through education and sympathy. The “sanctions” of morality are not metaphysical but human: the internal feeling of duty and the social consequences of behavior.
For Mill, moral rules are justified because they *promote happiness*; for Kant, they are justified because they *express rational necessity*.

(b) The Role of Consequences

Kant’s ethics is **non-teleological**: the rightness of an act depends on the motive of duty, not on the results it produces. Even a morally disastrous outcome does not taint a good will acting from duty.

Mill’s ethics is **teleological and consequentialist**: the rightness of an act is determined entirely by its consequences for happiness. The good (happiness) defines the right (those acts that maximize it).

Thus, Kant’s morality is grounded in *intention* and *universality*, whereas Mill’s is grounded in *outcomes* and *well-being*.

(c) The Role of Reason and Desire

Kant’s moral agent is a rational being who rises above inclination. Desire is a source of heteronomy — dependence on contingent motives. Acting morally means acting **from reason alone**.

Mill’s moral agent, in contrast, does not reject desire but **refines** it. Happiness, as the ultimate good, is a felt end rooted in human psychology. Reason’s role is to **organize and extend** this natural desire — to show that our own happiness is inseparable from the happiness of others. Mill’s moral development does not suppress inclination but **cultivates sympathy** and identifies moral action with enlightened self-interest.

(d) Universality and Motivation

Both philosophers aim at universal principles, but they differ in how universality operates.

For Kant, the universality of the moral law lies in its *form*: every rational agent can, by reasoning, recognize the same categorical imperative.
For Mill, universality lies in the *content* of human nature: all people desire happiness, and social progress consists in realizing this common good collectively.

Where Kant appeals to the authority of reason, Mill appeals to the universality of human sentiment.

(e) Two Visions of Moral Progress

Kant imagines moral progress as the triumph of reason over inclination — a rational autonomy that gives the moral law to itself.
Mill, by contrast, sees moral progress as the **education of feeling**, the expansion of sympathy, and the growing realization that individual happiness depends on the happiness of all.

In short:

Aspect	Kant	Mill
Basis of morality	Pure reason	Human nature, experience
Criterion of rightness	Conformity to duty	Maximization of happiness
Motivation	Acting from duty	Acting from enlightened sympathy
Moral law	Categorical imperative	Greatest happiness principle
Value of consequences	Irrelevant to morality	Constitutive of morality

(f) Conclusion: Two Paths to Moral Universality

Kant and Mill share the same ambition: to find a rational, universally valid basis for ethics beyond personal preference or cultural habit. Yet they travel opposite routes. Kant's moral law is **formal, a priori, and unconditional**; Mill's principle of utility is **empirical, teleological, and grounded in human well-being**.

Mill's justification of utilitarianism thus marks a decisive turn from the metaphysical to the psychological, from duty grounded in reason to duty grounded in **human nature and experience**. He invites us to see morality not as obedience to an abstract law but as the rational pursuit of the good life for all — a moral ideal inseparable from the flourishing of human sympathy, liberty, and progress.

Mill's Legacy: From Utilitarian Justice to Liberal Equality

Mill's effort to reconcile justice with utilitarianism represents one of the most ambitious moments in modern moral philosophy. By grounding justice in social utility, he sought to preserve the moral authority of rights within a consequentialist framework. His insight was that rights are not metaphysical entities but **social instruments** essential to human happiness — the moral infrastructure of a flourishing community.

In *On Liberty*, Mill extended this idea into the political sphere. Liberty was not only a moral right but a **prerequisite for collective progress**: freedom of thought, expression, and individuality generates the intellectual and moral vitality on which social improvement depends. Thus, protecting liberty is not opposed to utility — it is the surest means to secure it in the long run.

This interpretation transforms utilitarianism into a **liberal and developmental ethic**. The pursuit of happiness is not a matter of immediate gratification but of cultivating the higher faculties — reason, creativity, empathy — that sustain both individual dignity and social progress. In this respect, Mill's utilitarianism becomes less a calculus of pleasure than a philosophy of human flourishing.

The resonance of this vision can be traced in later liberal theories. **John Rawls's** *A Theory of Justice* (1971) can be seen as both a continuation and a corrective to Mill. Rawls rejected utilitarian aggregation as insufficiently protective of individual rights, proposing instead that justice be defined by fairness — by principles that no rational agents would reject under conditions of equality. Yet the **moral intuition** underlying Rawls's position — that rights and liberties are essential to the welfare of all — echoes Mill's attempt to anchor justice within the broader aim of human happiness.

In this sense, Mill stands at a pivotal moment in the history of moral thought:

- He inherited Bentham's faith in rational calculation but transformed it into a humane moral psychology grounded in sympathy and liberty.
- He anticipated Rawls's commitment to rights and fairness while maintaining the consequentialist vision of the good as human happiness.

Mill's justification of utilitarianism thus achieves more than a defense of one ethical theory; it articulates a **comprehensive vision of moral progress** — one in which reason, feeling, and social justice are mutually sustaining. His utilitarianism, enriched by a liberal concern for individuality and dignity, remains one of the most enduring attempts to unite moral obligation with human well-being.

2.7 The Population Challenge to Utilitarianism: Parfit's Repugnant Conclusion

Utilitarianism, from Bentham and Mill onward, has offered one of the most powerful frameworks for moral reasoning: maximize happiness, minimize suffering, and thereby promote the greatest good for the greatest number. Yet when the principle of utility is extended beyond individual ethics and into questions of population size and future generations, it leads to unsettling paradoxes. The most famous of these is Derek Parfit's *Repugnant Conclusion*, a problem that exposes the difficulty of applying utilitarian logic to the ethics of creation, reproduction, and intergenerational justice. Closely related is Parfit's *Non-Identity Problem*, which challenges not only utilitarianism but all moral theories premised on the idea of duties toward identifiable persons. Together, these paradoxes reveal that the utilitarian calculus, when applied to temporally extended populations, risks collapsing under the weight of its own moral assumptions.

The Repugnant Conclusion

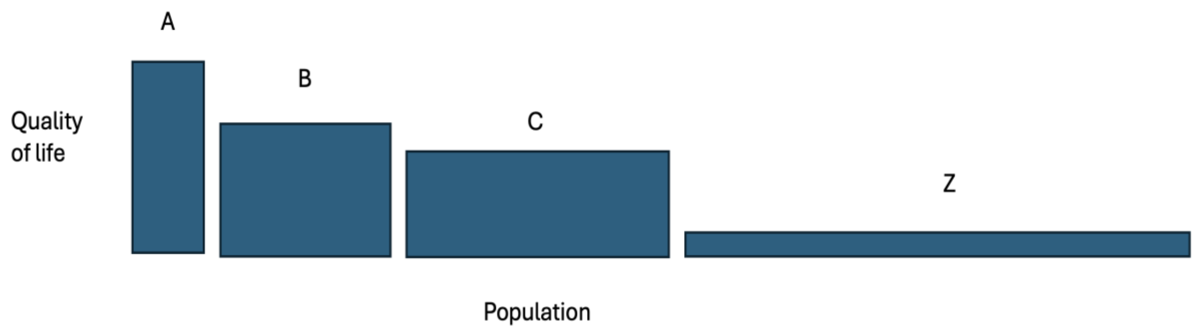


Fig. 2.1

Parfit's thought experiment, as shown in Fig. 2.1, begins with a world A, where a relatively small number of people enjoy very high-quality lives. We then imagine a world B, where there are more people, each living slightly less happily than in A. If the total sum of happiness in B is greater than in A, classical total utilitarianism dictates that B is a morally better world. Continuing this reasoning, we can imagine a long series of worlds—C, D, E, and so on—each with a larger population and slightly lower average happiness. Eventually we reach world Z, in which billions of people live lives that are just barely worth living. Yet because the total happiness in Z is higher than in A, utilitarianism tells us that Z is morally better than A.

This is the **Repugnant Conclusion**:

"For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better even though its members have lives barely worth living." *(Parfit, Reasons and Persons, 1984)*

The conclusion seems morally grotesque. It implies that a vast, mediocre population is better than a smaller, flourishing one. Yet it follows inexorably from the utilitarian commitment to maximizing total utility.

Average Utility and Its Limits

To escape this paradox, one might turn from *total* to *average utilitarianism*: instead of maximizing the sum of happiness, we maximize its average level across individuals. On this view, world A—where fewer people live exceptionally good lives—would be preferable to world Z, where happiness per person is minimal.

However, this solution leads to its own paradox. Suppose the world consists of a single "Paradise Island," with inhabitants enjoying happiness at a perfect score of 100/100. If we discover another island where people live good but slightly less happy lives (say, 90/100), the world's average happiness decreases, even though no one has become worse off and the total number of happy people has doubled. According to average utilitarianism, this discovery makes the world worse, which seems absurd. The addition of more happy people cannot plausibly make existence worse overall.

Thus, average utilitarianism, while avoiding the repugnant conclusion, produces what has been called the *sadistic implication*: it can prefer fewer people to more, even when all are happy. Both formulations—total and average—fail to reconcile our moral intuitions with utilitarian arithmetic.

The Non-Identity Problem

Parfit launched another, perhaps even more profound, challenge to utilitarianism through what he called the **Non-Identity Problem**. The argument begins from a simple but striking observation:

If a person is conceived at a different time, in different circumstances, they would in fact be a different person.

This insight has radical moral implications. Policies and actions that affect the timing and conditions of conception also affect *which* people come into existence. Hence, our present decisions shape not only the welfare of future individuals but also their very identities.

Consider two policies:

- **Policy A** involves the **depletion of natural resources**, leading to a higher standard of living for the current generation but leaving a degraded environment for future people.
- **Policy B** prioritizes **sustainability**, limiting consumption now to preserve resources and ecological balance for future generations.

If society adopts Policy A, many children will be conceived under current prosperous conditions; their descendants will live with fewer resources and lower environmental quality. If Policy B is chosen, future people will enjoy cleaner air and greater wealth—but crucially, they will not be the *same* individuals who would have existed under Policy A.

Here lies the paradox. If we choose the depletion policy, it seems we have not *harmed* anyone, since the people who will exist in the future would not have existed otherwise. The future generation that lives in polluted conditions cannot claim to have been wronged, for without the depletion policy, they would never have been born at all. They may regret their circumstances, but they can still rationally affirm that life is worth living.

From a **utilitarian standpoint**, this reasoning implies that no moral harm occurs: no existing person is made worse off. Therefore, utilitarianism seems to entail that we have *no duties* toward future generations—at least not in the person-regarding sense. The moral value of policies cannot be assessed in terms of harms or benefits to specific individuals, since the very identity of those individuals depends on the policies themselves.

The Utilitarian Response and Its Limits

A possible response is to abandon person-regarding reasoning altogether and assess policies purely by their **impersonal consequences**:

"We should not be concerned that under one policy certain people will exist and under another policy different people will exist. We should simply measure the amount of utility that each policy will bring to present and future people—whoever they are—and decide accordingly."

This impersonal approach aligns with utilitarianism's commitment to impartial aggregation, but it creates new difficulties when the **number of people** changes.

If we adopt **average utilitarianism**, we would favor policies that maximize average happiness, even if that means fewer people exist. Yet this conflicts with our moral intuition that a world containing more happy people, even if slightly less happy on average, is better than one with fewer happy people.

Conversely, if we adopt **total utilitarianism**, we would favor policies that increase the total number of people, regardless of how small their individual welfare may be—leading us straight back to Parfit's *Repugnant Conclusion*.

Thus, when population size is variable, utilitarianism lacks a stable criterion for moral evaluation. It oscillates between counterintuitive implications: either preferring smaller, happier populations (average utilitarianism) or vast, barely content ones (total utilitarianism).

Moreover, the **non-identity problem** exposes a deeper metaphysical tension: utilitarianism presupposes a fixed set of individuals whose utilities can be compared and aggregated. But in intergenerational contexts, this assumption collapses. There is no constant set of persons across possible futures—only potential populations with different members. This undermines the very basis of utilitarian calculation.

Beyond the Utilitarian Framework

Parfit himself did not abandon consequentialism. He proposed that moral reasoning should sometimes appeal to *impersonal principles*—judging the goodness of outcomes independently of who experiences them. Yet this move takes utilitarianism further from its intuitive foundation in the equal consideration of persons.

Later thinkers have tried to build more nuanced frameworks. **John Broome**, for instance, argued in *Weighing Goods*(1991) that population ethics must account for both fairness and total welfare, suggesting that moral value involves distributive as well as aggregative components. Similarly, **Blackorby, Bossert, and Donaldson** (2005) formalized how social choice theory cannot satisfy all plausible ethical conditions when population varies—mirroring Sen's impossibility results in welfare economics.

These efforts reveal that the problem is not simply one of calculation but of conceptual coherence. Once we allow population and identity to vary, the utilitarian model loses its grounding in a determinate moral community. It can no longer specify *who* counts, nor explain *why* existence itself should be morally valuable.

The challenge posed by population size and identity exposes the limits of utilitarianism's guiding ideal. When extended to questions of existence, reproduction, and future generations, the principle of maximizing happiness becomes indeterminate and self-undermining. Parfit's *Repugnant Conclusion* shows that total utilitarianism can justify worlds of minimal happiness, while the *Non-Identity Problem* suggests that utilitarianism cannot coherently express duties to future persons whose existence depends on present choices.

Both paradoxes stem from the same structural tension: utilitarianism's reduction of moral value to impersonal welfare aggregates. By treating well-being as a quantity to be maximized, it overlooks the moral significance of individuality, identity, and the complex temporal structure of human life.

Mill's own insights remain instructive here. His emphasis on higher pleasures, moral character, and human development gestures toward a richer conception of value—one that cannot be captured by mere sums or averages of happiness. The lesson of Parfit's paradoxes, then, is not that we should abandon concern for happiness, but that we must complement it with principles of dignity, justice, and respect for persons—across both time and identity.

2.8 The Contemporary Relevance of Utilitarianism: From Bentham to Singer

Two centuries after Jeremy Bentham formulated the *principle of utility*—the idea that right actions are those that promote the greatest happiness for the greatest number—utilitarianism remains one of the most influential frameworks in applied ethics. While early utilitarianism sought to reform institutions and legislation through a rational calculus of pleasure and pain, contemporary versions have extended its reach to the moral challenges of a globalized and interdependent world. Nowhere is this more evident than in the work of **Peter Singer**, whose synthesis of utilitarian reasoning and practical engagement has revitalized moral philosophy for the twenty-first century.

Bioethics: Life, Death, and the Value of Suffering

Utilitarianism entered bioethical debates through its insistence that moral decisions should aim to minimize suffering and maximize well-being, regardless of traditional taboos or metaphysical doctrines. Singer's *Practical Ethics* (1979; rev. 2011) applies this principle to questions such as euthanasia, assisted suicide, abortion, and the treatment of severely disabled infants.

In this framework, the moral status of an action depends not on conformity to rules or divine commands but on its **consequences for sentient beings**. Thus, prolonging life at all costs is not always morally required: if an individual's future holds only severe pain or loss of consciousness, ending life may reduce overall suffering. This reasoning challenges conventional sanctity-of-life ethics but aligns with Bentham's consequentialism and Mill's concern for the quality, rather than mere quantity, of life.

Singer's utilitarian approach has been both influential and controversial. Critics argue that it risks devaluing the lives of the vulnerable, while defenders note that it provides a rational and compassionate method for confronting medical dilemmas where moral intuitions conflict.

Animal Liberation and the Expansion of Moral Considerability

One of the most significant modern transformations of utilitarian thought concerns its extension beyond the human species. Bentham himself anticipated this shift when he wrote that the question is not whether animals can reason or talk, but whether they can suffer. Singer's *Animal Liberation* (1975) develops this insight into a systematic ethical theory.

Rejecting **speciesism**—the arbitrary preference for human interests over those of other sentient beings—Singer argues that animals' capacity to suffer gives them equal moral consideration. Consequently, practices such as **factory farming**, **animal testing**, and **industrial fishing** are morally indefensible, as they inflict immense suffering for comparatively trivial human benefits.

This utilitarian expansion has had profound cultural and institutional effects. It underpins the ethical rationale for **veganism and vegetarianism**, informs **animal welfare legislation**, and has shaped the discourse of environmental and food ethics. It transforms utilitarianism from a doctrine of human pleasure into a cosmopolitan ethics of sentient life.

Effective Altruism and Global Poverty

Singer's most far-reaching contemporary contribution is perhaps his leadership in the **effective altruism** movement. Building on the utilitarian logic of impartial benevolence, effective altruism seeks to maximize the positive impact of moral action by evaluating outcomes empirically—using evidence and reason to determine where resources can do the most good.

In *The Life You Can Save* (2009) and *The Most Good You Can Do* (2015), Singer argues that affluent individuals are morally obligated to donate a substantial portion of their income to highly effective charities addressing global poverty and preventable disease. The argument mirrors Bentham's universalist impulse: each person's suffering counts equally, whether they live next door or across the world.

The movement has inspired a new ethics of giving among younger generations, integrating data-driven philanthropy with moral philosophy. It also exemplifies **rule utilitarianism** in practice—establishing general principles and institutions that consistently maximize well-being over time rather than judging each act in isolation.

Environmental Ethics and Future Generations

Contemporary utilitarianism has also expanded its temporal scope, addressing the well-being of **future generations** and the moral implications of **climate change**. Singer and other consequentialists argue that environmental degradation—deforestation, species extinction, pollution, and global warming—produces vast and often irreversible harms to both humans and non-humans.

From a utilitarian standpoint, mitigating climate change is an ethical imperative: the suffering of future persons and animals counts no less than that of the present. This intergenerational concern re-engages the themes raised by Parfit's population ethics and the non-identity problem, showing that utilitarian reasoning, though challenged by such paradoxes, remains central to debates on sustainability and justice.

Singer thus connects environmental ethics with the same principle that motivated Bentham's social reforms: the impartial concern for all beings capable of happiness or suffering. The utilitarian vision becomes planetary in scope, urging not only reduced consumption and carbon emissions but a transformation of moral consciousness toward global responsibility.

Critiques of Singer's Utilitarianism

Despite its enduring influence, Singer's utilitarianism has faced significant criticism on both philosophical and moral grounds. **Deontological and rights-based theorists** argue that utilitarianism's exclusive focus on outcomes risks justifying morally objectionable acts—such as sacrificing an innocent person—if they produce a net increase in utility. From this perspective, utilitarian reasoning lacks respect for **individual rights and dignity**, treating persons as mere instruments of collective welfare.

Virtue ethicists contend that utilitarianism neglects the moral significance of character, motivation, and moral emotion. They argue that ethical life cannot be reduced to the calculation of outcomes, and that empathy, integrity, and moral growth are constitutive of the good, not merely instrumental to it.

In **bioethics**, Singer's views on euthanasia and infanticide have been accused of undermining the intrinsic worth of human life, especially that of the disabled or dependent. Critics such as Leon Kass and Michael Sandel claim that a purely utilitarian calculus fails to capture the moral meaning of care, dignity, and vulnerability.

Moreover, **communitarian and feminist theorists** have questioned the abstraction of utilitarian reasoning, noting that moral life is embedded in relationships, histories, and contexts that resist quantification. They argue that the pursuit of global impartiality may obscure local forms of solidarity and responsibility.

Nonetheless, even many critics concede that Singer's work has forced ethical discourse to confront uncomfortable questions with intellectual honesty. His arguments have expanded public awareness of suffering—human and non-human alike—and encouraged greater consistency between moral belief and everyday practice.

Conclusion: The Enduring Legacy of Utilitarian Reasoning

Two hundred years after Bentham's *Introduction to the Principles of Morals and Legislation* (1789), utilitarianism remains a living, adaptive tradition rather than a closed system. Its language of **consequences, welfare, and impartiality** continues to guide debates in medicine, animal rights, economics, and environmental policy.

Through Peter Singer and related movements, utilitarianism has evolved from a theory of social reform into a comprehensive **ethics of global responsibility**—one that evaluates choices by their real impact on the happiness and suffering of all sentient life, present and future.

While its critics highlight the risks of moral calculation and the neglect of rights or dignity, its defenders note that few ethical frameworks have done more to **reduce suffering, promote equality, and expand the circle of moral concern**. In this sense, modern utilitarianism preserves Bentham's original ambition: to make morality a matter not of authority or custom, but of reason and compassion applied to the shared conditions of life on earth.

BOX # 2.4: The Value of Life: A Utilitarian Perspective from Epicurus to Peter Singer

Few moral questions are as enduring and unsettling as that of life's value and the ethics of death. Philosophers, physicians, and policymakers continually confront situations where life must be weighed—against pain, against other lives, or against competing goods such as dignity and autonomy. While the idea of "weighing lives" strikes many as morally abhorrent, utilitarianism insists that such reasoning is unavoidable if we are to act rationally and compassionately.

The utilitarian tradition, from Bentham and Mill to contemporary thinkers such as Peter Singer, regards moral value as rooted in the balance of pleasure and pain, happiness and suffering. This quantitative, consequence-based framework contrasts sharply with the classical view of **Epicurus**, for whom death is nothing to fear because it cannot be experienced. The confrontation between these two traditions—Epicurean tranquility and utilitarian concern for the deprivation of future happiness—illuminates the continuing debate on the value of life and the moral permissibility of ending it.

Epicurus on the Meaninglessness of Death

Epicurus, writing in the *Letter to Menoeceus*, sought to liberate humans from the fear of death, which he saw as one of the principal sources of anxiety and unhappiness. His central claim is elegantly simple:

"Death is nothing to us; for when we exist, death is not present, and when death is present, we do not exist."

For Epicurus, the good is defined as pleasure—the absence of pain and disturbance (*ataraxia*). Death, being the total cessation of sensation, cannot be experienced and therefore cannot be painful or harmful. Fearing it is irrational because it concerns a state in which we no longer exist as subjects of experience.

This reasoning leads to two key implications:

1. Death is **not a harm** to the person who dies, because harm presupposes a subject who experiences it.
2. The wise person should cultivate indifference to mortality, focusing instead on achieving a life of tranquility and measured pleasure.

Epicurus's position is therapeutic and existential: it offers psychological liberation rather than a moral theory of value. Yet, it stands in tension with any ethical framework—like utilitarianism—that evaluates the goodness of states of affairs by their consequences for sentient beings, including the potential loss of future well-being.

The Utilitarian View: Life as a Container of Value

Utilitarianism reverses the Epicurean reasoning. While Epicurus denies that death can harm us, **utilitarians define the value of life by its potential for positive experience**. To die is to be deprived of those experiences, and thus death constitutes a harm—not because it is felt, but because it **prevents future happiness**.

This is often called the **deprivationist account of death**:

- Death is bad not as an experience but as a loss of valuable experiences one could have had.

- The harm of death lies in truncating the flow of pleasures, relationships, and satisfactions that make life good.

From this point of view, the value of life can be represented as the total quantity of happiness it contains. Doctors, for instance, implicitly employ this reasoning when comparing medical treatments: a short but high-quality life (Treatment A) may be preferable to a long but miserable one (Treatment B). The area under the "curve" of happiness over time—the *integral of well-being*—captures the life's total value.

Similarly, in triage situations, physicians may prioritize saving a younger person over an older one because the former has more potential years of happiness ahead. While emotionally distressing, this logic reflects the utilitarian imperative to maximize total welfare rather than respect life as intrinsically sacred.

For utilitarianism, therefore, **life is valuable not merely because it exists, but because it can be good**—because it can generate pleasure, love, understanding, and flourishing. A life devoid of such goods may lose its claim to be preserved at all costs

Death, Harm, and the Deprivationist Challenge

The **deprivationist critique of Epicurus**, advanced by philosophers such as Thomas Nagel and Derek Parfit, reinforces the utilitarian position. Nagel argues that the mere absence of experience does not negate harm: one can be harmed by deprivation even if one does not feel it. For example, if someone is betrayed without knowing it, their interests are still violated.

Analogously, death deprives a person of future goods they would otherwise have enjoyed, and thus counts as a harm, even if not experienced. This view aligns naturally with utilitarianism's consequentialist structure: what matters morally are the states of the world—the total distribution of happiness—not the subject's awareness of them.

From this perspective, prolonging life has moral value to the extent that it **increases the sum of positive experiences**. Conversely, ending life prematurely is bad insofar as it reduces the potential for future well-being—unless, crucially, continued existence would yield more suffering than happiness.

Utilitarianism and the Ethics of Euthanasia

It is precisely at this boundary—where life becomes more painful than pleasurable—that utilitarianism provides a distinctive rationale for euthanasia. **Peter Singer** has been the most influential modern philosopher defending the permissibility, and sometimes the duty, to end life under conditions of incurable suffering.

In *Practical Ethics* (1979; 2011), Singer argues that the value of life depends on its **quality of consciousness**, not merely its biological continuity. The moral justification for euthanasia derives from the **principle of minimizing suffering** and **respecting autonomous choice**.

- **Voluntary euthanasia**, when a competent person chooses to die to escape unbearable suffering, can increase overall utility by relieving pain and respecting individual preference.
- **Non-voluntary euthanasia**, in cases where a patient cannot decide (e.g., persistent vegetative state or severe infant disability), may also be justified if the life in question holds no prospect of net positive experience.

Singer rejects the notion that human life possesses intrinsic or sacred value apart from its capacity for sentient experience. This view is continuous with Bentham's original principle: "The question is not, Can they reason? nor, Can they talk? but, Can they suffer?" The same reasoning applies to humans who have lost the capacity for meaningful experience as to animals capable of suffering.

For Singer, to preserve a life that is dominated by agony or devoid of awareness may be morally worse than allowing death, since it increases total suffering without corresponding benefit. In contrast, to end such a life compassionately may express a deeper moral concern: the reduction of avoidable suffering for all sentient beings.

Objections and Ethical Tensions

Singer's utilitarian defense of euthanasia has provoked intense controversy. Critics argue that it undermines the inherent dignity of human life and risks legitimizing the killing of the disabled or elderly. From a **Kantian** or **rights-based** standpoint, treating life as a mere bearer of utility ignores its moral inviolability.

Others question the epistemic basis of utilitarian judgment: how can we reliably calculate the net happiness or suffering of continued existence, especially across different forms of consciousness? Moreover, by focusing exclusively on outcomes, utilitarianism may neglect relational and symbolic aspects of care—the values of compassion, respect, and solidarity that transcend utility.

Nonetheless, defenders maintain that Singer's position captures the ethical reality faced daily in medical contexts: decisions about life support, palliative care, and assisted dying inevitably require trade-offs between the quantity and quality of life. Utilitarianism at least offers a consistent and transparent criterion for making such decisions, one rooted in the moral significance of suffering rather than dogmatic prohibitions.

Conclusion

The confrontation between **Epicurus** and **utilitarianism** reveals two radically different attitudes toward death and the value of life. For Epicurus, death is nothing to us—a natural limit that should neither be feared nor morally weighted. For utilitarianism, by contrast, life derives its worth from the happiness it embodies and enables; death is a harm because it extinguishes that potential.

In modern bioethics, this utilitarian insight has become indispensable. Physicians, legislators, and ethicists must make decisions that balance lives, pain, and well-being. Through Peter Singer's work, utilitarianism continues to inform debates about euthanasia, end-of-life care, and the moral responsibilities that accompany technological medicine.

Ultimately, utilitarianism does not diminish the value of life—it seeks to **clarify it**. Life matters because it is the site of experience, joy, and consciousness; when those vanish irretrievably or are overwhelmed by suffering, the moral imperative may shift from prolonging existence to relieving pain. Two centuries after Bentham, this remains utilitarianism's enduring ethical legacy: a rational compassion grounded in the measurable realities of sentient life.

Chapter 3

LIBERTARIANISN

3.1 Introduction

In the moral and political debates of the twenty-first century, **libertarianism** continues to exert a powerful influence on how we understand freedom, justice, and the role of the state. Rooted in Enlightenment individualism and the moral primacy of self-determination, libertarianism offers a distinct vision of the good society—one in which individuals are free to pursue their own ends without coercion, and in which social order emerges from voluntary cooperation rather than centralized design. Yet as the global world confronts challenges such as **climate change, migration, global poverty, and persistent inequality**, libertarianism's insistence on minimal state intervention and moral self-responsibility raises pressing questions about its ethical adequacy and practical coherence.

At its philosophical core, libertarianism is grounded in the principle of **self-ownership**: each individual has an exclusive right to control their own body, labor, and legitimately acquired possessions. From this moral axiom flow the familiar libertarian rights to private property, freedom of contract, and resistance to coercion. But the moral psychology underlying these rights has deep affinities with **rational egoism** and **ethical egoism**, traditions that hold that rational agents ought to act in their own long-term interest and that moral value is rooted in individual well-being rather than collective welfare. For thinkers like **Ayn Rand**, rational egoism becomes a moral virtue: self-interest, when guided by reason and respect for others' autonomy, leads not to moral corruption but to human flourishing. In this sense, libertarianism can be understood as a social ethic that translates individual egoism into a political principle—freedom as the condition under which each can responsibly pursue their own good.

Different strands of libertarianism interpret this moral foundation in divergent ways. **Classical liberals** such as **Friedrich Hayek** and **Milton Friedman** defend markets not merely as efficient economic systems but as ethical mechanisms that respect individual choice and disperse knowledge through spontaneous order. **Anarcho-capitalists** like **Murray Rothbard** push this logic to its limit, claiming that all essential functions of the state—from policing to adjudication—can be provided through voluntary exchange. **Robert Nozick**, by contrast, offers a deontological justification for the **minimal state**, which exists solely to protect individuals from force, theft, and fraud. Beyond these, **left-libertarians** such as **Hillel Steiner** and **Philippe Van Parijs** argue respectively, that while individuals own themselves, the world's natural resources belong equally to all, thus requiring a fair distribution of their value and that it should be adopted a global **basic income** for all.

When viewed in the context of today's **global challenges**, libertarianism presents both promise and tension. On the one hand, its defense of voluntary cooperation, decentralized decision-making, and entrepreneurial initiative resonates with global efforts to empower individuals and reduce dependency on paternalistic states. Its ethos of **self-responsibility** can inspire innovation, civic autonomy, and ethical restraint. On the other hand, libertarianism's skepticism toward collective action and redistributive policies appears ill-suited to address problems that transcend individual boundaries—such as **climate change**, **mass migration**, and **world poverty**.

In environmental ethics, for instance, libertarianism confronts a paradox: if pollution and carbon emissions harm others' property or well-being, they seem to violate the non-aggression principle. Yet global environmental degradation often involves diffuse, cumulative harms that no single agent can easily internalize or redress. Some libertarian thinkers have sought to extend property-rights reasoning to environmental protection—through tradable pollution rights, private conservation, or compensation schemes—but such market-based solutions may fall short of the cooperative scale that the climate crisis demands.

In the ethics of **migration**, libertarian principles lead to striking conclusions. Since individuals have a right to move freely and engage in voluntary exchange, many libertarians support open borders as a moral corollary of economic freedom. Yet others argue that unrestricted migration may threaten the integrity of voluntary associations or the rule of law on which liberty depends. The debate reveals a deep tension between universalist and communitarian readings of libertarian ethics: **is freedom a cosmopolitan right or a contract among members of a given polity?**

Finally, libertarianism's approach to **global inequality and poverty** exposes one of its central moral controversies. From a utilitarian or Rawlsian perspective, justice requires positive duties to assist the least advantaged. Libertarians, however, deny that moral obligation can entail coercive redistribution. For **Nozick**, taxing individuals to fund welfare programs amounts to a form of forced labor. Nevertheless, **contemporary libertarians** inspired by global justice debates have explored voluntary and market-based means of reducing poverty—microfinance, open trade, and technological empowerment—arguing that freedom itself is the most powerful engine of development.

Libertarianism's enduring appeal lies in its moral clarity: it begins with the inviolable dignity of the individual and builds outward, constructing a vision of social order that seeks justice through the protection of liberty. Yet as this chapter will explore, the global age compels libertarian thought to confront new questions: can a moral theory grounded in individual rights and rational self-interest respond adequately to collective existential risks? Can it reconcile personal freedom with the moral interdependence of a shared planet? And might a renewed libertarian ethics—sensitive to environmental limits, human vulnerability, and global justice—offer not only a defense of liberty but a redefinition of it for a connected and fragile world?

3.2 Murray Rothbard: Libertarianism, Natural Rights, and the Critique of the State

Murray Rothbard (1926–1995) was an American heterodox economist of the **Austrian School** and a political theorist whose writings played a seminal role in shaping the **modern American libertarian movement** and the doctrine of **anarcho-capitalism**. His work combines Austrian economic theory, Lockean natural rights, and a radical critique of state power.

He viewed the state not as a neutral guarantor of order but as an **institution of systematic coercion**. He famously described it as *"the organization of robbery systematized and writ large."* In his view, every function performed by the state—defense, policing, infrastructure, or education—could be provided more efficiently and morally by voluntary, private arrangements within a free market. This conviction led him to oppose not only socialism and welfare statism but even **minimal-state libertarianism** (as in Nozick's "night-watchman" model), arguing that any state authority necessarily violates individual rights.

Rothbard was a fierce critic of **central banking** and the **fractional-reserve banking system**, which he saw as fraudulent and inflationary. He rejected both **Keynesian economics**, which he dismissed as shallow opportunism, and **mainstream classical and neoclassical economics**, including figures like **Adam Smith** and **David Ricardo**, for failing to recognize the moral and epistemological foundations of economic freedom. Even among libertarian economists, he diverged sharply—accusing **Milton Friedman** of "statism" for endorsing monetary policies that preserved the Federal Reserve's role.

Rothbard's economic system ideal was a **pure laissez-faire capitalism**, rooted in voluntary exchange and private property, completely free of government interference through regulation, subsidies, or monetary manipulation.

While many defenders of capitalism, including Ludwig von Mises, grounded their arguments in **utilitarianism**—that markets maximize social welfare—Rothbard rejected this consequentialist foundation. He accused Mises of inconsistency for defending liberty on utilitarian grounds while claiming to uphold value-free economics. For Rothbard, utilitarianism was too permissive: it justified state interventions whenever they might increase aggregate welfare.

Instead, Rothbard advanced a **natural rights** foundation for libertarianism, borrowing heavily from **John Locke's** theory of property and self-ownership. He saw human rights as objective and derived from human nature—part of a universal **natural law** tradition. Because humans are rational, self-directed beings, any coercive interference in their person or property violates their nature and moral order.

The Non-Aggression Axiom

At the heart of Rothbard's philosophy lies the **Non-Aggression Axiom**:

"No man or group of men may aggress against the person or property of anyone else."

Aggression—defined as the initiation of physical force or fraud—is illegitimate, regardless of motive or consequence. From this axiom flows the moral defense of private property, voluntary exchange, and free contract. It also entails absolute opposition to taxation (*"taxation is robbery"*), conscription (*"slavery on a mass scale"*), and war (*"mass murder"*).

Property and Self-Ownership

Rothbard's ethics begin with **self-ownership**: every person has an exclusive right to control their own body and life. Denying this principle leads to either (a) slavery, where one class owns another, or (b) communal ownership of all individuals by everyone, which is both impossible and self-destructive.

Building on Locke's **labor-mixing** theory, Rothbard argued that individuals acquire legitimate ownership over external resources by transforming them through labor ("homesteading"). He dismissed "communal" or "collective" claims to property as incoherent and prone to oligarchic domination. His interpretation included the **Lockean proviso**—that appropriation is legitimate so long as it does not deprive others of access to unowned resources—but he applied it minimally, prioritizing the individual's productive transformation of nature (see BOX # 1).

Society and the Individual

Rothbard denied that "society" exists as an independent moral entity capable of acting or owning. Society, he insisted, is merely a **collection of individuals**. To speak of "social rights" or "collective interests" is, therefore, a category mistake—a rhetorical cover for coercion. In this, Rothbard echoed both **individualist anarchism** and **Margaret Thatcher's** dictum: "There is no such thing as society; there are individual men and women, and there are families."

War, Civil Liberties, and the State as Aggressor

Rothbard viewed **war** as the ultimate manifestation of state aggression: *"War is the health of the state."* He believed militarism perpetuates centralized power, taxation, and economic distortion. His opposition to war was so radical that he condemned all conflicts except two: the **American Revolution** and the **Confederate War of Secession**, which he interpreted (controversially) as defenses of self-determination against centralized tyranny.

Consistent with his anti-statism, he championed **civil liberties**: freedom of speech, press, assembly, and the right to engage in "victimless crimes" such as drug use, prostitution, and pornography. To him, any consensual act among adults was outside the moral or legal jurisdiction of the state.

Libertarianism, Rational Egoism, and Ethical Egoism

Rothbard's libertarianism also resonates with **rational and ethical egoism**. Like Ayn Rand, he believed that individuals ought to act in rational pursuit of their own interests, provided they do not aggress against others. The moral legitimacy of the free market, in this view, lies in its alignment with human nature as self-directed and self-interested. Rational egoism—the idea that reason and self-interest coincide—is compatible with his natural-law ethics: each individual, as a rational being, must be free to choose and act on their own values. However, unlike Rand's objectivist ethics, Rothbard's foundation was not metaphysical egoism but **moral natural law**. Egoistic self-assertion was justified insofar as it respected the identical rights of others.

Property Rights as Human Rights

For Rothbard, **property rights and human rights are identical**. For example, some liberal (in opposition to libertarians) stand for personal liberties such as freedom to speech, press, assemble etc., but they are ready to compromise or to deny rights to property. According to Rothbard, the two (property rights and human rights) are inextricably intertwined. They stand or fall together. For example, the liberal socialist advocates government ownership of all means of production while standing for the human right of freedom of speech or press. But how is the right of free press to be exercised if the government owns all the newsprints and all the printing shops. The same is true for the right of free speech if the government owns all the assembly halls and therefore allocates those halls as it sees fit.

Egalitarianism and Its Discontents

Rothbard was an uncompromising critic of **egalitarianism**, which he saw as an unnatural and destructive ideology. Equality, he wrote, "is not in the natural order of things," and efforts to impose it—beyond equality before the law—inevitably lead to coercion, resentment, and economic decay. His opposition extended to the **civil rights movement**, which he viewed as an instance of government overreach in private association and contract.

3.3 Murray Rothbard: Libertarian Natural Law and the Limits of the State

Murray Rothbard's thought represents one of the most radical expressions of libertarian philosophy in the twentieth century. Building upon the Austrian School of economics and John Locke's natural rights theory, he developed a comprehensive ethical, political, and economic system grounded in the principles of **self-ownership**, **private property**, and **voluntary exchange**.

Rothbard rejected all forms of state coercion, regarding the state as "the organization of robbery systematized and writ large." His **Non-Aggression Axiom** holds that no person or group may initiate force against another's person or property. From this principle follows his defense of laissez-faire capitalism and his rejection of taxation, conscription, and welfare redistribution.

Unlike **utilitarianism**, which justifies institutions according to their capacity to maximize aggregate welfare, Rothbard grounded libertarianism in **natural law**. For him, rights are not contingent on consequences or utility but derive from human nature itself: since human beings are rational and purposive agents, they must be free to act according to their reasoned choices. This view distinguishes him not only from utilitarian economists such as Jeremy Bentham and John Stuart Mill, but also from his mentor **Ludwig von Mises**, whose utilitarian arguments for market efficiency Rothbard rejected as morally insufficient.

Libertarianism, Utilitarianism, and Contractarianism

Rothbard's ethics stand in stark contrast to the two dominant traditions of modern liberal thought: **utilitarianism** and **contractarianism**.

Where utilitarianism justifies moral and political rules by their capacity to produce the greatest happiness for the greatest number, Rothbard saw this as a dangerous and relativistic foundation. Because utilitarian reasoning evaluates actions by outcomes, it can justify coercive policies if they appear to maximize welfare—for example, taxation for public goods, compulsory education, or redistributive welfare. To Rothbard, such reasoning abandons moral principle for expediency and undermines the inviolability of individual rights.

Similarly, while **John Rawls's contractarianism** shares with libertarianism a concern for individual liberty, it departs radically in its egalitarian implications. Rawls's **difference principle**, which permits inequalities only if they benefit the least advantaged, legitimizes redistributive measures that Rothbard would condemn as violations of property rights. For Rothbard, justice is not a pattern to be achieved but a condition defined by the **absence of aggression**. He would therefore regard Rawlsian distributive justice as a moral error: an attempt to impose an artificial social order through coercion.

Even **Robert Nozick's minimal state libertarianism**, though closer to Rothbard's view, stops short of full anarchism. While Nozick defended a "night-watchman" state limited to the protection of rights, Rothbard insisted that even this minimal state violates the Non-Aggression Principle, since taxation and monopoly enforcement are forms of coercion. For him, the only legitimate social order is one based entirely on **voluntary associations** and **private enforcement of contracts**—a fully **anarcho-capitalist society**.

3.4 Environment, Migration, and Global Inequality

Rothbard's libertarianism offers a radical reinterpretation of many contemporary social and political challenges. His commitment to **natural rights**, **private property**, and

voluntary exchange as absolute moral principles leads him to sharply diverge from both utilitarian reformers and social-democratic egalitarians.

a) Environment and Property Rights

Rothbard's environmental thought emerges from his conviction that all social ills stem from violations of property rights. Environmental degradation, in his view, results from **unowned or collectively owned resources**, which invite overuse and neglect. The remedy is not regulation or international agreements but the **extension of private ownership**to all resources, including air, water, and land.

Pollution would thus constitute an **act of aggression** under the Non-Aggression Axiom, subject to private legal redress. Victims could sue polluters for damages, as property owners whose rights have been violated. Rothbard rejected environmental regulation and taxation as unjust forms of coercion, arguing that free-market mechanisms and private stewardship provide more effective incentives for conservation.

This position contrasts sharply with both **utilitarian environmental ethics**, which justify regulation by reference to aggregate welfare (e.g., cost–benefit analysis), and **Rawlsian approaches**, which emphasize intergenerational justice and the fair distribution of environmental goods. Rothbard's view privileges the **rights of present individuals** over the collective claims of future generations or society as a whole.

b) Migration and the Right to Association

On migration, Rothbard's framework yields an ambivalent position. Consistent with his principles of liberty and voluntary exchange, he acknowledged the right of individuals to move and contract freely. Yet he argued that in a society based entirely on private property, there would be no "public space" to cross freely—entry would depend on the consent of property owners.

Hence, while **utilitarian and egalitarian theories** might defend open borders on humanitarian or welfare grounds, Rothbard's anarcho-capitalism would base migration entirely on **voluntary association**. Movement across territories becomes a matter of property rights, not state policy. Critics have observed that this logic could legitimize exclusion and privatized discrimination, but Rothbard regarded such outcomes as morally permissible within the framework of voluntary relations.

c) Global Poverty and Inequality

Rothbard rejected all forms of redistributive justice, both within and between nations. In his view, poverty is not an injustice but a **natural consequence** of human diversity and free competition. Attempts to alleviate poverty through taxation, welfare, or foreign aid constitute coercion and violate property rights.

Against global egalitarian theories such as **Peter Singer's utilitarianism** or **Thomas Pogge's cosmopolitan contractarianism**, Rothbard argued that moral obligations cannot be imposed beyond voluntary consent. The only legitimate form of aid is voluntary charity, not institutional redistribution. For Rothbard, global justice consists not in equal outcomes but in **universal respect for the Non-Aggression Principle**—that is, in ending coercion, protectionism, and state intervention that distort free exchange.

This position sharply diverges from utilitarian and Rawlsian global ethics, which see distributive justice and collective responsibility as necessary correctives to market inequalities. For Rothbard, the pursuit of equality leads inevitably to tyranny; liberty, not equality, is the highest moral and political value.

At the psychological and ethical level, Rothbard's theory resonates with **rational and ethical egoism**—the view that acting in one's rational self-interest is both moral and natural. Individuals are not bound by altruistic duties but by the moral requirement to respect others' autonomy. Cooperation, charity, or environmental care are justified only when **freely chosen** as expressions of one's rational will, not imposed by law or collective morality.

This egoist foundation gives Rothbard's libertarianism both its coherence and its limits: it upholds a vision of justice as **non-interference** and morality as **self-direction**, but it also struggles to account for collective goods, shared risks, and structural inequalities in a globally interdependent world.

Conclusion: Rothbard and the Global Ethics of Freedom

Murray Rothbard's libertarianism extends classical liberalism to its philosophical limit. By grounding justice in the absolute prohibition of coercion, he rejects both the **utilitarian** appeal to welfare and the **contractarian** concern for fairness. His system upholds liberty as the supreme moral value, viewing the state as an obstacle to justice rather than its guarantor.

In the context of the **global world**, however, this radical individualism encounters new tensions. Environmental degradation, mass migration, and global poverty reveal dimensions of interdependence that transcend voluntary exchange. While Rothbard's framework highlights the dangers of coercive paternalism, it also exposes the ethical challenges of sustaining freedom in a world of shared vulnerability.

Two centuries after Bentham and two decades after Rawls, libertarianism continues to ask a fundamental question: **Can justice be achieved without coercion?** Rothbard's answer—an unyielding defense of natural liberty—remains one of the most provocative contributions to the philosophy of the global age.

3.5 Robert Nozick: Self-Ownership, Justice, and the Minimal State

Robert Nozick (1938–2002) was born in Brooklyn, New York, into a Jewish immigrant family. In his youth, he joined Norman Thomas's Socialist Party, attracted by its egalitarian ideals. Over time, however, his intellectual trajectory shifted dramatically toward libertarianism, influenced by thinkers such as Friedrich Hayek and the anarcho-capitalist economist Murray Rothbard.

Nozick's early socialist sympathies evolved into a deep concern for individual autonomy and property rights, leading him to reject state redistribution and paternalism. His 1974 book *Anarchy, State, and Utopia* became one of the most influential and controversial works in twentieth-century political philosophy, offering a powerful response to John Rawls's *A Theory of Justice* (1971) and reshaping the debate over freedom, justice, and the legitimate scope of government.

Entitlement Theory and the Principle of Self-Ownership

At the core of Nozick's political philosophy lies his **entitlement theory of justice**, grounded in the idea that individuals are inviolable ends in themselves. Drawing on the Lockean tradition, Nozick argues that people own themselves and, through their labor, may acquire previously unowned resources. Land and other natural resources do not belong to men, however if men have originally mixed their labor with them they are entitled to own those resources as far as the ***Lockean Proviso*** is respected (see Box #1). Justice in holdings, therefore, depends on three principles:

1. **Justice in acquisition**, the legitimate initial appropriation of unowned resources;
2. **Justice in transfer**, voluntary exchanges or gifts between consenting individuals; and
3. **Rectification of injustice**, which addresses holdings acquired through force, theft, or fraud.

If these principles are respected, any resulting distribution—no matter how unequal—is just **(just-preserving principle)**. A just society, then, is not defined by a pattern of outcomes but by the fairness of the processes through which property is obtained and exchanged. Redistribution for the sake of equality violates this entitlement, as it takes from some without their consent and thereby infringes their rights.

BOX # 3.1: Locke's Original Proviso

In *Two Treatises of Government* (1690), Locke argued that individuals can claim ownership over parts of the natural world by "mixing their labor" with them. However, this process is subject to a moral constraint: one may appropriate from nature **only insofar as there remains "enough and as good left in common for others."** This clause—the **Lockean proviso**—ensures that private ownership does not deprive others of access to the means necessary for survival and flourishing. In Locke's moral universe, property is legitimate only when it does not make others worse off than they would have been in a world without private ownership.

Nozick's Reformulation: From Equality to Non-Worsening

Nozick accepts Locke's basic insight but reinterprets the proviso in less egalitarian and more libertarian terms. For Nozick, a person's appropriation of an unowned resource is just **if and only if it does not worsen the position of others** who are thereby excluded from using it. He writes that:

"A process of appropriation is just if it does not worsen the situation of others who are no longer at liberty to use the thing."

This reformulation weakens Locke's original egalitarian thrust. Where Locke's proviso required that enough resources remain for others, Nozick's version requires only that others are not made *worse off* than they would have been in a pre-property state of nature. The moral baseline is thus not equality, but the *absence of harm*.

The Water Hole Example

Nozick illustrates this through the case of a water hole in a desert. If an individual fences off the water source but continues to allow others to draw from it on reasonable terms, no one is made worse off, and the appropriation is legitimate. However, if the owner excludes others or charges exorbitant prices that threaten their survival, the proviso is violated. The legitimacy of property, therefore, hinges not on abstract fairness but on whether others' opportunities for well-being are reduced below the natural baseline.

The Role of the Proviso in the Entitlement Theory

The Lockean proviso serves as a **moral constraint** on the first principle of Nozick's Entitlement Theory—the **principle of just acquisition**.
According to this theory, a distribution of goods is just if it arises from three legitimate processes:

1. **Just acquisition** – the initial appropriation of unowned resources in a manner consistent with the proviso;
2. **Just transfer** – the voluntary exchange of holdings through trade, gift, or contract;
3. **Rectification** – compensation or correction for past injustices in acquisition or transfer.

As long as these conditions are met, the resulting distribution—no matter how unequal—is just. In Nozick's words, "from each as they choose, to each as they are chosen."

Justice as Historical, Not End-State

Nozick's insistence on the proviso reflects his broader view that justice is **historical**, not **patterned**. A distribution cannot be judged fair or unfair merely by looking at its current pattern (for instance, how much wealth each person has), but only by understanding the historical processes that produced it. If property was acquired and transferred without violating anyone's rights, then any resulting inequalities are morally permissible—even if they appear unjust by egalitarian or utilitarian standards. Thus, the Lockean proviso functions as a safeguard ensuring that original acquisition does not constitute an act of aggression or deprivation.

Minimal Constraint and Market Society

In practice, Nozick treats the proviso as a **minimal constraint**, one that rarely restricts ownership in developed market economies. The growth of productivity and voluntary exchange, he argues, tends to make everyone better off than they would have been in a pre-property state of nature. Consequently, the proviso is satisfied whenever the free market raises general welfare—even if it produces vast inequalities. For Nozick, this justifies both the strong protection of property rights and the rejection of redistributive policies.

Critiques and Implications

Critics contend that Nozick's interpretation of the Lockean proviso dilutes Locke's original moral intent. Whereas Locke sought to balance liberty with fairness, Nozick reduces justice to non-interference and historical legitimacy. His version permits extreme inequalities as long as the worst-off are not literally worse off than before. This, critics argue, overlooks structural injustices and the cumulative effects of economic power. Yet for Nozick and libertarians more broadly, the proviso's minimalism is its virtue: it prevents property from becoming morally arbitrary while preserving the core libertarian commitment to self-ownership and voluntary exchange.

Historical Justice and the Critique of Redistribution

For Nozick, justice is *historical*: it depends on how things came to be, not on whether the resulting pattern fits an ideal model. Redistribution policies—such as progressive taxation or welfare programs—are unjust because they take the fruits of one person's labor to benefit another without consent.

He famously compared such compulsory transfers to forced labor: "Taxation of earnings from labor is on a par with forced labor." By taking part of what a person earns, the state effectively claims partial ownership over that person's time and abilities.

However, Nozick acknowledges that not all existing property arrangements are just, since history is full of coercion and expropriation. His rectification principle suggests that past injustices—such as colonization, slavery, or land theft—must be addressed before a libertarian system can claim moral legitimacy. Yet he provides few concrete guidelines for how rectification should occur, leaving this as an open and contentious issue, especially in global contexts marked by historical inequality.

The Emergence of the Minimal State

Nozick begins from a Lockean *state of nature* in which individuals possess natural rights to life, liberty, and property. These rights are moral constraints on the actions of others and cannot be overridden for the sake of collective welfare.

He then develops a thought experiment to show how a **minimal state**—limited to protecting individuals against force, theft, and fraud, and enforcing contracts—could arise without violating anyone's rights. In a stateless society, people would hire *private protection agencies* to safeguard their rights. Over time, one agency might become dominant, forming an *ultraminimal state*.

To prevent conflict, this dominant agency would gradually extend protection to all individuals, including those unable to afford it, creating a *minimal state* that protects everyone's basic rights but performs no redistributive or paternalistic functions. Importantly, this state does not emerge by coercion but through an "invisible-hand process" of voluntary interactions—making it morally justified.

Nozick's moral foundation owes much to Immanuel Kant. He interprets Kant's principle that human beings must be treated as *ends in themselves* to mean that individuals have inviolable moral boundaries. Using someone merely as a means to promote social welfare—through forced labor, taxation, or redistribution—fails to respect their inherent dignity and autonomy.

Thus, the minimal state is not merely an economic ideal but a moral one: it ensures that individuals remain sovereign over their lives and talents, free from coercive impositions justified in the name of collective good.

Critique of Utilitarianism: The Experience Machine and the Utility Monster

Nozick's defense of individual rights also entails a powerful critique of utilitarianism. He argues that utilitarianism's focus on maximizing happiness or welfare can justify violating individual rights if doing so increases overall utility. To expose this flaw, he introduces two famous thought experiments:

- **The Experience Machine:** Imagine a device that could provide you with any pleasurable experience you desire, indistinguishable from reality. Would you plug in? Most people would not, Nozick argues, because we value *authenticity*—being in touch with reality and shaping our own lives—not merely pleasure. This shows that human well-being involves more than utility maximization.
- **The Utility Monster:** Imagine a being who gains vastly more pleasure from resources than anyone else. Utilitarianism would require giving everything to this being, sacrificing everyone else. The absurdity of this outcome reveals utilitarianism's inability to respect individual rights and boundaries.

For Nozick, these arguments demonstrate that rights serve as moral "side constraints" that protect individuals from being sacrificed for others' benefit. Rights are not to be weighed against utility—they are absolute moral boundaries (see also BOX # 3 for the Omela's parable).

BOX # 3.2: The Omelas' parable.

Ursula K. Le Guin's celebrated short story *The Ones Who Walk Away from Omelas* (1973) offers one of the most striking literary critiques of utilitarian moral reasoning. Set in the idyllic city of Omelas, the story describes a society whose happiness and prosperity depend entirely on the perpetual suffering of a single child, confined in misery and neglect. Every citizen knows that their collective joy—the harmony, abundance, and moral beauty of their city—rests upon the torment of this innocent being. While most accept this arrangement as a regrettable necessity, some, unable to reconcile such happiness with justice, choose to leave Omelas and walk away into an unknown future.

Le Guin's fictional city embodies the logic of **classical utilitarianism**, which holds that the moral rightness of an action or institution depends on its ability to maximize overall happiness or utility. In Omelas, this principle is realized to perfection: the suffering of one ensures the happiness of thousands. If moral value consists only in the balance of pleasure over pain, then Omelas is an exemplary society. Yet the story's emotional force lies precisely in showing that such a society, however prosperous, is **morally intolerable**.

The moral cost of Omelas reveals what many critics of utilitarianism—especially those within libertarian and deontological traditions—regard as its deepest flaw: its readiness to sacrifice the rights or well-being of the few for the happiness of the many. The child's suffering becomes a vivid symbol of what is obscured by aggregate moral reasoning: the **inviolable moral status of individuals**.

Omelas's story echoes Robert Nozick's "Utility Monster", a being that derives vastly more pleasure from resources than anyone else. According to strict utilitarian logic, we should therefore give everything to the monster, since this maximizes total happiness. Like the child in Omelas, the Utility Monster dramatizes the **tyranny of aggregation**—the idea that moral justification can rest solely on maximizing total utility, regardless of how happiness is distributed or who suffers in the process.

Both Le Guin and Nozick confront readers with the same moral intuition: that **individual rights and dignity set boundaries** to what can be done in the name of collective welfare. No amount of overall happiness can justify the deliberate harm or instrumentalization of a person.

Nozick's "Experience Machine" further challenges utilitarian hedonism by asking whether we would choose to live in a machine that guarantees perpetual pleasure but disconnects us from reality. Most people, he suggests, would refuse—because we value authenticity, autonomy, and moral integrity more than mere pleasurable experience.

When read alongside *Omelas*, the Experience Machine underscores another weakness of utilitarianism: it reduces all values to subjective states of pleasure or satisfaction, ignoring the importance of truth, justice, and moral respect. The citizens of Omelas may feel deep joy, but their happiness rests on a lie—a denial of the moral reality of the child's suffering. Their world, like the machine's simulated reality, is a form of moral self-deception.

For libertarian thought, the story of Omelas illuminates the centrality of **individual rights** as constraints on collective calculation. Nozick's minimal state, founded on the inviolability of self-ownership, would reject the very premise of Omelas: that one person may be sacrificed, however involuntarily, for the benefit of others. In this sense, Le Guin's parable illustrates why libertarian and deontological ethics insist that individuals are **ends in themselves**, not instruments of utility.

At the same time, those "who walk away from Omelas" represent a form of moral integrity that transcends utilitarian calculus. They choose uncertainty and loss over a happiness built on injustice. Their departure evokes the libertarian ideal of **moral autonomy**—the refusal to participate in systems that violate the rights and dignity of others, even when doing so appears socially beneficial.

Le Guin's *Omelas* challenges readers to ask a question that echoes across moral philosophy: **Is happiness still good when it is unjustly bought?** In confronting this paradox, the story becomes more than a critique of utilitarianism—it is an invitation to reconsider the foundations of justice itself.

From a libertarian perspective, the lesson of Omelas is clear: the moral legitimacy of a society does not depend on the sum of its pleasures but on the respect it shows for the freedom and inviolability of each individual. A just order, whether national or global, must ensure that no one's suffering becomes the hidden price of others' prosperity. In a world facing persistent inequality, environmental degradation, and structural exploitation, *Omelas* remains a timeless warning: a society that measures good solely by outcomes risks forgetting the humanity of those it leaves behind.

BOX # 3.3: The Wilt Chamberlain Example

Nozick begins by noting that many theories of justice in distribution rely on a **pattern** — for example:

- everyone should have **equal shares**; or
- people should have resources distributed according to **moral merit**, **need**, or **contribution**.

He calls these "patterned principles" because they prescribe that justice depends on some particular pattern or end-state of holdings across society.

To show the instability of such patterned distributions, Nozick constructs a thought experiment:

- Suppose we begin with a perfectly just distribution of resources, **D_1**, according to your favorite egalitarian pattern (say, everyone has $100).
- Now, imagine the basketball player **Wilt Chamberlain** signs a contract stating that each fan who attends his games voluntarily gives him 25 cents.
- A million fans willingly attend and pay, so Chamberlain earns $250,000 more than anyone else.

We now have a new distribution, **D_2**, where Wilt is far richer than others.

Nozick's point is: if **D_1** was just, and all the transfers from **D_1 to D_2** were **voluntary**, then **D_2** must also be just — because justice in holdings depends on **how** people acquire and transfer resources, not on **what the resulting pattern looks like**.

Therefore, to restore the original pattern (equality, say), the state would have to **interfere** — by **forcibly taking away** some of Wilt's earnings.
That, Nozick argues, **violates individual rights**, especially the right to **self-ownership** (the idea that people own themselves and, by extension, their labor and what they earn through voluntary exchange).

The Wilt Chamberlain example thus functions as a **critique of redistributive justice**:

- Redistribution is unjust because it **disrespects individual entitlements**—it treats people as means to achieve a collective pattern rather than as ends in themselves.
- Any **patterned** principle of justice is inherently unstable: voluntary exchanges will always disrupt it, requiring **continuous state coercion** to maintain.
- Justice, according to Nozick, should be **historical** (based on how holdings came about) rather than **end-state**(based on what the distribution looks like).

In short:

"Liberty upsets patterns."
Maintaining a preferred distributive pattern means limiting people's liberty to make voluntary exchanges.

This argument rests on Nozick's **libertarian** view of rights:

- People own themselves.
- Any interference with their property or earnings (beyond minimal state functions like protecting against force and fraud) is unjust.
- Redistribution, even for egalitarian aims, amounts to **partial slavery**, because it forces some to work for the benefit of others.

Libertarianism and Global Ethics: The Minimal State in a Globalized World

In the context of globalization, Nozick's libertarian theory raises profound questions about the moral scope of justice beyond national borders. His framework, grounded in self-ownership and voluntary exchange, offers a powerful defense of negative rights—freedom from coercion and interference—but leaves limited space for positive obligations such as global redistribution or collective environmental responsibility.

From a **global justice** perspective, Nozick's rejection of patterned or outcome-based principles implies that inequalities between nations are not inherently unjust, provided they result from voluntary interactions rather than coercion or theft. This view contrasts sharply with cosmopolitan theories such as those of Peter Singer or Thomas Pogge, who argue that affluent societies have moral duties to alleviate extreme poverty worldwide. For Nozick, compulsory taxation to fund global aid programs or climate adaptation schemes would constitute an infringement of individual property rights. Moral responsibility, in his view, arises from *voluntary charity*, not enforced solidarity.

In the debate on **migration**, however, Nozick's principles yield a paradox. If individuals own themselves and have the right to exchange their labor and resources freely, then, in principle, they should also have the right to move across borders in search of better opportunities. Yet many libertarian theorists defend strong property-based rights of exclusion, viewing state borders as extensions of collective ownership. This tension reveals an unresolved issue in libertarian thought: whether self-ownership supports *open borders* as an expression of individual liberty, or *closed borders* as an exercise of collective property rights.

Environmental challenges further test the limits of Nozickian libertarianism. Climate change, resource depletion, and biodiversity loss are collective-action problems that cannot easily be addressed through individual contracts. The libertarian emphasis on property rights may encourage environmental stewardship when resources are privately owned and protected against overuse—the so-called "tragedy of the commons" solution. However, when ecological harms are diffuse and transboundary, the minimal state lacks the authority to coordinate large-scale mitigation or adaptation efforts. Nozick's framework thus struggles to address *intergenerational justice* and global environmental interdependence, where harm is often unintentional but widely distributed.

At the same time, Nozick's vision of a society based on **voluntary cooperation, respect for individual autonomy, and limits on coercive power** remains an important normative counterpoint in global ethics. It challenges paternalistic or technocratic approaches to development, reminding policymakers that moral progress cannot be achieved by violating individual rights in the name of collective welfare. In this sense, libertarianism continues to play a vital role in the philosophical dialogue about freedom, responsibility, and justice in a globalized world.

3.6 Left-Libertarianism: Self-Ownership and Equality of Resources

Left-libertarianism represents an important contemporary development within the broader libertarian tradition. It seeks to combine the **libertarian commitment to individual self-ownership** with the **egalitarian concern for a fair distribution of the world's natural resources**. Thinkers such as **Peter Vallentyne**, **Hillel Steiner**, and **Philippe Van Parijs** have been central in articulating and defending this position.

Between Self-Ownership and Equality

Left-libertarians share with right-libertarians, such as **Robert Nozick**, the conviction that each person has a **universal and inviolable right of self-ownership**. This means that individuals own themselves, their bodies, and the fruits of their labor. Such ownership implies extensive negative rights: the right to be free from coercion, the right to control one's own actions, and the right to retain the material benefits that one's work produces.

However, unlike Nozick, left-libertarians argue that **natural resources**—land, minerals, air, water, and the like—are not rightfully owned by any individual in an absolute sense. They belong originally to everyone in common. Therefore, while one may appropriate and use these resources, such appropriation must respect **Locke's famous proviso**: it is just only if "enough and as good" is left for others. In the modern interpretation offered by Vallentyne, Steiner, and Van Parijs, this proviso entails that **individuals have no moral right to an unequal share of the world's natural endowment** without compensating others.

The Egalitarian Implications of Locke's Proviso

This egalitarian reading of Locke leads left-libertarians to propose systems that guarantee **equal access or compensation** for the private use of natural resources. One central policy implication is the idea of a **universal basic income (UBI)**—a guaranteed minimum payment to all individuals, financed by taxing the value of natural resources, land rents, or other collectively owned assets.

For **Philippe Van Parijs**, in *Real Freedom for All* (1995), this form of redistribution is not a violation of liberty but a **precondition of genuine freedom**. He distinguishes between merely formal freedom—the absence of interference—and **"real freedom"**, which requires the material means to pursue one's life plans. A universal basic income, funded through common ownership of external resources, secures for all individuals a fair starting point to exercise their liberty meaningfully.

BOX # 3.4: The idea of Universal Basic Income

Philippe Van Parijs is a Belgian political philosopher whose most influential work, *Real Freedom for All: What (if Anything) Can Justify Capitalism?* (1995), provides a normative defense of **Universal Basic Income (UBI)**.

A **Universal Basic Income** means:

- a **regular, unconditional cash payment** to every individual,
- **universal** (everyone receives it, regardless of income, work, or status),
- **individual** (not household-based),
- **unconditional** (not tied to employment or behavioral requirements).

Van Parijs's argument starts from a simple premise:

What matters for justice is not merely formal freedom (having legal rights), but **real freedom** — the actual ability to act on one's life plans.

He defines *real freedom* as **"the effective capacity to do whatever one might want to do."** This links to a broader egalitarian idea: everyone should have **the greatest possible real freedom** compatible with others having the same.

Van Parijs argues that **UBI best realizes "real freedom for all"** because it provides people with the **material means** to make genuine choices — including the choice **not to work**.

Key points in the argument:

a. Freedom vs. Constraint

In a capitalist economy, people without resources must sell their labor to survive.
This dependence limits their real freedom — they are "formally" free to do anything, but in practice they must take whatever job they can get.

A **basic income** breaks that constraint by guaranteeing a secure, unconditional income floor.
It enables people to:

- refuse exploitative work,
- pursue education or caregiving,
- engage in creative or civic projects,
- or simply take time off without coercion.

Thus, UBI expands **real** rather than merely **nominal** freedom.

Unlike welfare programs that are **conditional** (means-tested, work-tested), UBI is **unconditional** and **universal**.
This ensures:

- **no stigma** or bureaucratic control,

- **no disincentive to work** (since the income is not withdrawn as one earns more),
- and **equal treatment** of all citizens as free agents.

For Van Parijs, such unconditionality expresses a deep form of **equal respect** and **reciprocity**.

Van Parijs calls his view **"real libertarianism."**
He agrees with libertarians (like Nozick) that freedom is central, but insists that *freedom must be real and equally distributed*, not merely formal.

He argues that natural resources, technology, and social capital — the "common inheritance of humankind" — are collectively produced.
Therefore, all individuals have a **moral claim** to share in that wealth, even if they do not contribute directly through labor.

In this sense, UBI is a **dividend** from shared social resources, ensuring that everyone benefits from the collective product of civilization.

Van Parijs's UBI can be read as a **direct answer to Nozick's Wilt Chamberlain critique**:

Nozick	Van Parijs
Redistribution violates self-ownership; it takes from the entitled.	Basic income honors self-ownership by maximizing real freedom for all, funded from common resources (not others' labor).
Liberty upsets patterns — voluntary exchanges destroy equality.	Liberty without basic security is hollow; UBI ensures genuine, not merely formal, liberty.
Justice is historical (based on voluntary transfers).	Justice must ensure fair access to opportunities and resources that are collectively created.

So, while Nozick defends *freedom from interference*, Van Parijs defends *freedom to act* — freedom **as capacity**, not just **as non-interference**.

Van Parijs's practical formula for justice is:

Each person should receive the highest sustainable unconditional basic income compatible with maintaining economic viability.

That is, UBI should be as high as possible without undermining productive cooperation or creating unsustainable incentives.

This makes his theory both **normatively egalitarian** and **economically pragmatic**.

Critics have raised several objections:

- **Exploitation concern**: Is it fair that some live off others' labor?
  Van Parijs replies that all wealth depends on shared social and natural capital, so no one is a "free rider."
- **Feasibility concern**: Can economies sustain such payments without inflation or reduced productivity?

He admits limits but argues that technological progress and capital accumulation make UBI increasingly realistic.

- **Motivation concern**: Would people still work?
  He claims many would, since work offers social and intrinsic rewards beyond survival.

Contrast with Nozick's Libertarianism

In contrast, **Robert Nozick's** right-libertarianism allows vast inequalities in wealth and opportunity, provided they arise from voluntary exchanges and just initial acquisitions. Nozick interprets Locke's proviso weakly: so long as private appropriation does not make others worse off than they would be in a world without private property, inequality is permissible.

Left-libertarians reject this minimal interpretation. They argue that Nozick's view effectively legitimizes arbitrary inequalities rooted in luck and birth, rather than in genuine individual merit. For Vallentyne and Steiner, justice requires that individuals share equally in the value of external goods, even if they remain fully entitled to the fruits of their labor. The **self-ownership principle** is therefore combined with a **right to an equal share of the value of natural resources**, producing a synthesis of libertarian and egalitarian ethics.

Left-libertarianism has important implications for **global justice**. If natural resources belong equally to all human beings, then global inequalities in wealth and opportunity—often based on the unequal geographic distribution of those resources—require rectification. A global version of the universal basic income, financed through taxes on the use of the planet's common resources (such as atmospheric carbon sinks, seabed minerals, or global trade routes), would express this principle at the international level.

Ultimately, left-libertarians maintain that **freedom and equality are not opposed but complementary**. True liberty, they argue, requires not only that individuals be free from coercion but also that they enjoy fair access to the world's common inheritance. By combining the **libertarian right to self-ownership** with the **egalitarian right to equal resources**, thinkers such as Vallentyne, Steiner, and Van Parijs offer a compelling vision of a society where **every person possesses both ownership over themselves and a fair share in the world that sustains them**.

Concluding Reflection

In the context of an increasingly interconnected and unequal world, **left-libertarianism** provides a valuable framework for rethinking justice beyond national boundaries. It preserves the libertarian respect for personal autonomy and voluntary exchange while addressing issues of **global poverty, migration, and environmental degradation** that stem from the unequal distribution of the earth's resources. By grounding equality in the shared ownership of the planet's natural wealth, left-libertarianism invites reflection on how freedom can be made meaningful for all—**not by restricting liberty, but by universalizing its conditions**.

3.7 Libertarianism and Future Generations

Libertarians generally believe that individuals have no *coercive* obligation to future generations but may have a *moral*or *voluntary* responsibility to consider their well-being. Their stance derives from core principles of self-ownership, non-aggression, and individual freedom. Any duty imposed by the state to secure benefits for future persons—such as through taxation, redistribution, or regulation—would, from the libertarian perspective, constitute an unjustified infringement on the liberty and property rights of present individuals.

1. No Forced Obligations
Libertarians oppose government-imposed duties on the present generation to benefit future ones. Obligations should arise only through voluntary agreements, not through coercive mechanisms. Hence, policies that restrict present freedom in the name of future welfare—such as heavy environmental regulations or intergenerational transfers through public debt—are seen as paternalistic and ethically problematic.

2. Voluntary Responsibility
While rejecting coercion, many libertarians acknowledge that individuals and communities may freely choose to act responsibly toward future generations. Innovation, conservation, and prudent resource use are viewed as rational actions when they align with self-interest, market incentives, or personal moral values. Voluntary institutions, philanthropic efforts, and technological solutions are thus considered legitimate and effective ways to secure long-term sustainability without violating individual liberty.

3. Property Rights and Market Solutions
Libertarian theory maintains that strong property rights and free markets provide the most reliable framework for long-term environmental and intergenerational stewardship. Private ownership encourages resource holders to preserve and enhance the value of their assets over time—whether forests, fisheries, or farmland. Through mechanisms such as inheritance, investment, and market exchange, individuals have incentives to manage resources sustainably and ensure that their property retains value for future generations.

4. Opposition to Government Intervention
Libertarians are generally skeptical of policies justified on behalf of future generations when they rely on government coercion. State-imposed regulations, subsidies, or environmental mandates are viewed as distortions that generate inefficiencies and infringe on individual autonomy. Similarly, accumulating public debt for future-oriented projects is seen as a form of intergenerational exploitation, transferring costs to people who did not consent to them. From a libertarian standpoint, the most ethical approach is to allow free individuals to coordinate through voluntary exchange and innovation rather than through political compulsion.

5. Ethical Variations and Philosophical Disagreement
Within libertarian thought, ethical interpretations differ. Some theorists argue that we cannot have obligations toward individuals who do not yet exist, since moral duties presuppose identifiable persons. Others, adopting a more virtue-ethical or prudential view, maintain that a society valuing liberty must also cultivate foresight and responsibility. By fostering respect for property, knowledge, and technological progress, free societies can promote the conditions that enable future prosperity without coercion.

6. Comparison with Rawls and Utilitarianism
In contrast to libertarian skepticism toward coercive obligations, both Rawlsian contractarianism and utilitarianism extend moral concern explicitly to future generations. John Rawls, in *A Theory of Justice*, formulates the **"just savings principle,"** which requires each generation to preserve sufficient resources to sustain just institutions for those yet to come. This principle, while not prescribing equal sacrifice across generations, establishes a collective duty of stewardship grounded in fairness rather than voluntary choice. Utilitarianism, by contrast, treats the well-being of future persons as part of the aggregate calculus of happiness: policies are justified insofar as they maximize total welfare across time. From a libertarian standpoint, both approaches risk subordinating individual rights to abstract collective goals. The libertarian insists that moral consideration for future people must remain within the bounds of present consent and property rights, not imposed by state authority. Yet, in practice, libertarian advocacy of voluntary conservation and innovation may coincide with utilitarian or Rawlsian outcomes, illustrating how differing moral premises can converge in promoting sustainability and intergenerational prudence.

Libertarianism, Future Generations, and Global Ethics

From a global perspective, libertarian thought presents a distinctive and often controversial stance on intergenerational and transnational justice. In issues such as climate change, global poverty, or migration, libertarians generally reject coercive redistribution or state-enforced environmental mandates, emphasizing instead the creative and adaptive potential of free markets, technological innovation, and voluntary cooperation.

They argue that by securing individual rights, private property, and the freedom of exchange, societies naturally generate the incentives and resources necessary to address long-term challenges. Innovation, entrepreneurship, and decentralized decision-making are viewed as more effective and morally legitimate tools than centralized governmental control.

Yet this position also raises important philosophical questions. Can voluntary action and market incentives alone secure justice for those who do not yet exist? Does liberty entail moral responsibility beyond the present generation? These questions reveal both the strengths and the limits of libertarian ethics in the global age—its deep faith in human agency and voluntary cooperation, but also its uncertainty about how far duties can extend when the beneficiaries are future, and as yet nonexistent, persons.

3.8 Conclusion: Libertarianism in a Shared and Wounded World

Libertarianism begins from a compelling moral intuition: that each person is an inviolable owner of themselves, and that justice consists above all in protecting individuals from coercion. From Rothbard's anarcho-capitalism to Nozick's minimal state and the egalitarian refinements of left-libertarianism, this tradition offers a powerful reminder that political projects which disregard individual rights in the name of progress or welfare risk turning liberation into domination. In a world scarred by authoritarianism, colonialism, and technocratic hubris, libertarianism's insistence on consent, voluntary cooperation, and the dangers of paternalism remains an indispensable critical resource.

Yet when we move from the domestic sphere to the global stage—from the relatively bounded world of contracts and holdings to the vast, interdependent systems that structure climate, migration, and world poverty—the limitations of libertarianism become increasingly apparent. The core concepts of self-ownership, non-aggression, and historical entitlement were not designed to address problems that are diffuse, cumulative, and structurally embedded. Greenhouse gas emissions, for instance, do not neatly map onto identifiable acts of aggression by one agent against another; they emerge from billions of dispersed decisions, interacting over decades, across borders, and across generations. A theory that recognizes only wrongful interference between present individuals struggles to capture the moral urgency of harms that are global in scope and largely borne by the poor and the unborn.

Something similar occurs with migration and global poverty. Taken strictly, libertarian self-ownership and freedom of contract seem to support open borders and free movement; yet, once property rights are treated as absolute, they can equally be invoked to justify exclusion and privatized gatekeeping. The question of who may cross which borders cannot be resolved simply by appealing to individual rights in a world where access to opportunity is profoundly shaped by historical injustices, arbitrary frontiers, and collective decisions. Likewise, a framework that treats poverty as misfortune rather than injustice, and that rejects any coercive redistribution, proves normatively thin when confronted with massive, systematic global deprivation. Voluntary charity and market-led growth may indeed alleviate some suffering, but they offer no principled way of addressing structural disadvantages rooted in colonial history, unequal resource endowments, or global economic rules.

Left-libertarians attempt to repair some of these deficits by insisting on equal claims to the world's natural resources and by defending mechanisms such as a (global) basic income funded from common assets. This move acknowledges that self-ownership alone cannot deliver justice in a world where the "starting gate" is radically unequal. Yet even here, the theoretical vocabulary remains primarily individualistic: the focus is still on balancing separate claims to resources, rather than on transforming the collective institutions that generate vulnerability and domination. Problems like climate change, systemic exploitation in global supply chains, or unjust migration regimes require coordinated political action and shared responsibilities that are hard to describe purely in the language of individual entitlements and voluntary exchange.

In light of these tensions, the global age exposes both the enduring insights and the blind spots of libertarianism. It rightly warns against sacrificing persons to aggregates, and against using the rhetoric of justice to legitimize new forms of control. But its strong presumption against collective coercion makes it difficult to acknowledge that some of our deepest injustices are not reducible to discrete violations of property or contract; they are built into the background structures that shape who is free, who is mobile, who is secure, and who bears the costs of environmental and economic crises. In a shared and wounded world, justice cannot be only a matter of “not interfering” with those who already have; it must also ask what we owe to those whose life chances are systematically constrained by arrangements from which others benefit.

For a philosophy of the global world, libertarianism is thus best understood not as a complete ethics of justice, but as a vital, partial perspective. It sharpens our sensitivity to the dangers of coercion and the moral significance of individual agency. Yet to confront world poverty, migration, and climate change, we must supplement its insights with normative frameworks that can take seriously collective goods, structural power, and intergenerational responsibility. The challenge is not simply to choose between liberty and justice, but to rethink liberty itself as something that can only be fully realized within just and sustainable common institutions—institutions capable of protecting individuals and addressing the large-scale harms and dependencies that no one, acting alone, can remedy.

Chapter 4

RAWLS'S CONTRACTUALISM

4.1 Introduction

Contractualism is a philosophical term referring to a family of political theories within the **social contract tradition**. Since its early modern roots in the works of Hobbes, Locke, and Rousseau, contractualism has sought to explain how legitimate political order and principles of justice can emerge from the voluntary agreement of free and rational individuals.

In recent decades, social contract theories have expanded their scope beyond explaining political authority. They now also serve as a foundation for theories of **distributive justice**, addressing how economic and social goods should be fairly allocated within society.

Those who enter into a contract are typically conceived as **rational, self-interested agents**, whose identities and rights exist independently of any particular social arrangement. From this perspective, the contract is not imposed but freely accepted because it offers mutual benefits: it establishes rules that enable cooperation, protect individuals from harm, and promote security and prosperity. Thus, the **contractarian** may argue that it is rational to accept certain restrictions on one's freedom in some areas—such as taxation, law enforcement, or property regulation—in order to gain the broader advantages of living in a just and stable society.

One of the distinctive features of contractualism is that, although it begins from **individualistic premises**, it ultimately seeks to justify **social cooperation**. The social contract, in this sense, is not merely a pragmatic arrangement but a moral framework that explains why individuals have reason to endorse shared norms of justice.

Within this broad family of theories, two main **traditions** can be distinguished:

1. **The tradition of mutual advantage**, associated with Hobbes, Hume, game theory, and David Gauthier.
2. **The tradition of impartiality**, associated with the Stoics, Kant, and John Rawls.

Both traditions share three fundamental assumptions:
a) questions of justice arise where there are **conflicts of interest** among individuals; and

b) **justice consists in what all persons could, in principle, agree to**, if they reasoned rationally and fairly.

However, they diverge in how they interpret the nature and motivation of justice.

The **mutual advantage** tradition views justice as a matter of **necessity and compromise**. Individuals cooperate not out of moral duty, but because cooperation produces better outcomes than pure self-interest would allow. As Gauthier puts it, *justice is rational prudence in contexts where cooperation enables us to achieve what we desire*. The state of affairs prior to any agreement—what philosophers call the "initial position"—is therefore crucial. The benefits that each individual can expect from the contract depend on their **original bargaining power** and resources, which means that existing inequalities often shape the terms of agreement.

By contrast, the **impartiality** tradition, inspired by Kant and developed by Rawls, defines justice not as a compromise but as an **impartial standpoint**. To act justly means to adopt the **moral point of view**—to see oneself in the position of others and to defend principles that could be accepted by anyone, regardless of their particular circumstances or interests. In this view, the motive for justice is not self-interest but **moral reasoning**: one has reason to accept a principle of justice because it is fair, not because it serves one's advantage. Justice, therefore, is independent of bargaining power. The just agreement is the one that all could endorse **as free and equal persons**, without knowing their place in society.

In sum, contractualism offers a powerful framework for understanding justice as a product of rational agreement. Whether grounded in mutual advantage or in impartial fairness, it provides a bridge between **individual freedom** and **social cooperation**, showing how self-governing individuals can come together to build a just political and moral order.

4.2 John Rawls

John Rawls was born in Baltimore, Maryland, in 1921, the second of five sons. His father was a respected lawyer, and his mother was an active member of the League of Women Voters, an organization devoted to expanding democratic participation and civic responsibility. Rawls graduated from Princeton University, where he began developing his lifelong interest in moral and political philosophy.

During World War II, he served as an infantryman in the Pacific, fighting in New Guinea and the Philippines. The horrors he witnessed during the war deeply affected him, leading him to lose his Christian faith. Promoted to sergeant, he was later demoted for refusing to carry out an order he considered morally wrong—a decision that reveals much about his character and moral convictions. The atomic bombing of Hiroshima also left a lasting impression on him. After the war, Rawls participated in rallies in Washington against the use of nuclear weapons and, later, in the 1960s, he voiced doubts about the legitimacy of the United States' involvement in Vietnam.

In 1971, Rawls published *A Theory of Justice*, a work that would become one of the most influential texts in twentieth-century political philosophy. Since its publication, it has inspired an ongoing and worldwide debate on the principles of justice, equality, and fairness. The book quickly became a cornerstone of university courses in political and social philosophy, and it reshaped the way philosophers think about the moral foundations of liberal democracy.

Over the following decades, Rawls refined and clarified many of the ideas first introduced in *A Theory of Justice* without ever abandoning its central theoretical structure. In this chapter, we will not follow a strictly historical or philological approach but instead focus on the essential ideas of his theory, while taking into account how his thought evolved over time.

When Rawls first developed his *Theory of Justice*, he did not explicitly consider whether his conception could be accepted by people who hold different comprehensive doctrines or views of the good life—whether religious, philosophical, or moral. At that stage, his theory could be read as a comprehensive moral doctrine itself, offering a complete account of how to live and what values to pursue.

However, Rawls later came to recognize that in a modern democratic society, it is impossible to expect all citizens to agree on a single conception of the good. The coexistence of diverse religious, moral, and philosophical beliefs is not a sign of social failure but an inherent and permanent feature of democratic pluralism. What matters, then, is not that everyone shares the same vision of the good, but that they can agree on a common political conception of justice that regulates the basic structure of society in a fair way.

In *Political Liberalism* (1993), Rawls revised his earlier position, distinguishing between comprehensive doctrines and political conceptions of justice. He argued that the conception presented in *A Theory of Justice* should not be seen as a comprehensive moral doctrine but as a *political* conception—one that citizens with different worldviews can endorse for different reasons. The stability of a democratic society, Rawls now claimed, depends on whether its conception of justice can be reasonably accepted by a broad and enduring majority of citizens who nonetheless hold diverse beliefs.

Thus, the task of political philosophy, for Rawls, is to find principles of justice that can command an *overlapping consensus*—a shared agreement among free and equal citizens who, despite their moral or religious differences, recognize the fairness of the same political framework. In this way, Rawls sought to reconcile liberty, equality, and pluralism within a stable and just democratic order.

The publication of *A Theory of Justice* in 1971 profoundly reshaped modern political philosophy. Rawls's conception of *justice as fairness* restored moral reasoning to the center of political thought, challenging the utilitarian assumption that justice could be reduced to the maximization of collective welfare. By grounding justice in the ideas of fairness, equality, and social cooperation among free and equal persons, Rawls offered a new way of understanding the moral foundations of democratic institutions.

Rawls's work also set the stage for some of the most important philosophical debates of the late twentieth century. Robert Nozick's *Anarchy, State, and Utopia* (1974) responded directly to Rawls, arguing from a libertarian standpoint that any redistributive principle of justice violates individual rights to self-ownership and voluntary exchange. Amartya Sen, on the other hand, challenged Rawls's emphasis on ideal institutional design, proposing instead a *capabilities approach* that evaluates justice in terms of people's real freedoms and opportunities to achieve well-being.

Beyond these critiques, Rawls's framework inspired new directions in global political theory. Thinkers such as Thomas Pogge, Charles Beitz, and Martha Nussbaum extended or questioned Rawls's principles at the international level, asking whether fairness should apply not only within national societies but also among peoples in an unequal global order. These global readings of Rawls have helped shape contemporary discussions on poverty, inequality, human rights, and climate justice—issues that define our interconnected world.

In this sense, Rawls's thought continues to provide a moral compass for reimagining justice beyond borders. His vision of a fair system of cooperation among free and equal individuals remains a cornerstone for any serious philosophical inquiry into how liberty and equality can coexist in a just global order.

4.3 The Five Fundamental Ideas of Rawls's Theory

The practical goal of Rawls's theory of justice is to identify a moral and philosophical foundation for democratic institutions—one capable of guiding the definition of rights, liberties, and equality. The first key idea in pursuit of this goal is to regard political society as a fair system of social cooperation, one that endures over time and across generations. Social cooperation, for Rawls, does not result from a centralized plan or the command of a social planner; rather, it depends on public rules that are recognized and followed by citizens who understand themselves as free and equal participants. To cooperate means to accept terms of collaboration that are seen as fair, and to recognize that adherence to these terms yields reciprocal advantages for everyone involved.

Closely related to this is the idea of a **well-ordered society**—a community governed by a public conception of justice that is shared and acknowledged by all. In such a society, each person accepts the principles of justice with the awareness that others do the same. A society that fails to meet this condition, Rawls argues, is fundamentally unstable, for it lacks the mutual trust required for lasting cooperation. His effort, therefore, is directed toward identifying a conception of justice capable of achieving consensus among citizens who adhere to diverse religious, moral, and philosophical doctrines. The plurality of "comprehensive doctrines" is, for Rawls, not a defect but a defining and permanent feature of modern democratic societies. If a society were to flatten itself into a single moral or religious creed, it would risk sliding into totalitarianism. The wager of Rawls's political liberalism lies precisely in the belief that there can exist a **political conception of justice** broad enough to serve as the moral basis for social unity within a pluralist democracy.

A further fundamental idea is that of the **basic structure of society**, the network of major political, economic, and social institutions that define the framework within which cooperation takes place. The basic structure determines how fundamental rights and duties are distributed and influences the life prospects and opportunities of individuals from birth. It includes the constitution, the judiciary, the laws of property, the economic system (such as the existence of a market), and other major social institutions. The political conception of justice regulates this structure by establishing the principles that guide it. These principles are not applied directly to each association or institution—such as churches, universities, or families—but they nevertheless constrain their actions. Thus, a church may excommunicate its members but not persecute them; a university may not discriminate against students; a family cannot neglect or harm its children. In this way, the principles of justice provide a moral framework for the entire social order.

The **original position** represents the fourth and perhaps most distinctive idea in Rawls's political philosophy. It serves as a conceptual device to identify fair principles of social cooperation. Rawls asks: what principles would free and rational individuals choose if they were to design the basic terms of their society without knowing their future position within it? To answer this, he imagines a hypothetical situation in which the parties find themselves behind a "veil of ignorance" that hides from them their social status, class, natural talents, ethnicity, or wealth. Under such conditions of impartiality, the principles agreed upon would be fair, because no one could shape them to their personal advantage. The original position thus functions as a rational procedure for deriving the principles that should regulate the basic structure of society, grounding justice as fairness in a model of equality and reciprocity.

The fifth and final idea concerns **persons as free and equal**. For Rawls, citizens must possess two "moral powers" that enable genuine social cooperation: first, the capacity for a sense of justice—that is, the ability to understand and apply the principles defining a fair social order—and second, the capacity to form, pursue, and revise their own conception of the good. These capacities belong to all individuals and form the basis of their moral personhood. The sense of justice develops through active participation in a just society, while the ability to conceive of the good reflects each person's moral freedom—the freedom to choose and, if necessary, to revise their own values and aims. Here Rawls's thought resonates with the Kantian tradition: as for Kant, human freedom expresses itself in obedience to moral laws one gives to oneself. Equality, in turn, derives from the fact that all individuals, despite their differences, share the same moral powers that qualify them to take part in social cooperation. Being free and equal persons, therefore, is not a metaphysical condition but a social and political fact, one that underpins the democratic ideal of citizens participating with equal rights and duties in the life of the community.

These five fundamental ideas—**social cooperation**, the **well-ordered society**, the **basic structure**, the **original position**, and **persons as free and equal**—form an integrated framework that defines Rawls's conception of justice. Each idea supports the others: social cooperation becomes stable when guided by a public conception of justice; this conception governs the basic structure of society; the fairness of its principles is guaranteed by the original position; and its legitimacy rests on the equal moral powers of free citizens.

Through this architecture, Rawls provides not only a theory of justice but also a philosophical justification for the moral stability of democratic institutions. It is from this conceptual foundation that he proceeds to formulate his two famous **principles of justice**, which specify how liberty and equality should be distributed within a well-ordered democratic society.

4.4 The Principles of Justice

Differences and inequalities in life prospects among individuals are, to a large extent, determined by factors that are beyond one's control or will. These factors include one's social class of origin, the natural talents a person happens to possess, the quality of education accessible to them, and, last but not least, sheer luck. The central question Rawls asks is: **what principles can make such differences legitimate and consistent with the idea of citizens as free and equal participants in a fair system of social cooperation?** In any society where citizens pursue their own aspirations and life plans, certain inequalities of wealth and status are inevitable. Yet, in Rawls's view, a society can develop in a just and orderly way only when its members perceive those inequalities as legitimate—namely, as the outcome of fair and commonly accepted rules.

Rawls argues that inequalities are acceptable only if they comply with two fundamental **principles of justice**:

1. **Each person has an equal right to the most extensive set of basic liberties compatible with similar liberties for others**; and
2. **Social and economic inequalities must satisfy two conditions**:
    - they must be attached to positions and offices open to all under conditions of fair equality of opportunity; and
    - they must be to the greatest benefit of the least advantaged members of society (the **difference principle**).

The first principle takes priority over the second, just as the principle of fair equality of opportunity takes precedence over the difference principle. This means that there can be no trade-offs between basic rights and liberties, on the one hand, and social or economic advantages, on the other. For example, it would be unacceptable to limit certain groups' fundamental liberties on the grounds that doing so might lead to greater economic efficiency or material welfare.

Importantly, Rawls's notion of *basic liberties* refers not to some abstract or purely formal conception of freedom but to a historically specific set of rights and freedoms secured through political and civil struggles in democratic societies. These liberties can be grouped into three main categories:

- **political rights** such as freedom of speech, association, and the right to vote;
- **personal or civil liberties** such as freedom of conscience, thought, and non-political association;
- and **protective rights** ensuring personal integrity, such as freedom from psychological or physical coercion, protection against arbitrary arrest, and access to fair and transparent judicial procedures.

With respect to **fair equality of opportunity**, Rawls insists that public offices and positions of responsibility must be open to all not merely in a formal sense, but in a *substantive* one: individuals with similar talents and abilities should have genuinely equal chances to attain them, regardless of their social background.

Rawls also notes a crucial difference between his two principles. The first principle typically forms part of constitutional provisions, while the principle of fair opportunity—and the difference principle—require specific interventions in the **basic structure** of society. In a market system, for example, to prevent excessive concentration of wealth and political power over time, corrective mechanisms such as inheritance taxation are necessary to preserve the conditions of fair opportunity.

For Rawls, a constitutional regime is stable only when there exists a shared commitment to the first principle across the political spectrum. As he puts it, *"It is this agreement that makes the government legitimate in its intention and the opposition loyal in its opposition. When both sides display mutual loyalty and recognize the agreement, a constitutional regime is assured."* (JaF, p. 49)

The second principle, by contrast, concerns redistributive justice and usually falls outside constitutional frameworks. It is often more contested and requires ongoing adjustments in the basic structure to be effective. Still, some version of the difference principle must always be present if social cooperation is to remain fair and if the least advantaged members of society are not to feel that their interests have been sacrificed to the benefit of the more fortunate.

Rawls emphasizes that the problem of distributive justice is **not** how to divide a fixed social product in a way that maximizes overall satisfaction or utility. Rather, it is a matter of defining fair rules for the basic structure that yield both an efficient and a just productive system. In other words, justice should be built *ex ante* into the rules of cooperation, not applied *ex post* to redistribute outcomes. Incentive schemes that reward the more talented or productive may be justified—*but only if they ultimately benefit the least advantaged*.

The **basic structure** thus becomes the primary focus of Rawls's political conception of justice. In his early work, Rawls had been concerned with the moral character and conduct of individuals. But he soon realized that the major moral questions of modern societies could only be addressed through an analysis of their institutions. Social injustices stem from institutional arrangements, and it is therefore institutions—not individuals—that must be reformed. Moreover, since social and economic systems evolve, justice must be actively maintained through continuous adjustment.

This marks a key departure from the classical liberal tradition rooted in John Locke. According to the **just-preserving principle**, if the initial distribution of wealth is just, any subsequent distribution remains just so long as it arises from voluntary exchanges and transfers. Rawls rejects this view. Even if the initial distribution were just, he argues, it would likely become distorted over time due to the accumulation of advantages and the effects of luck, corruption, market fluctuations, or systemic crises. These dynamics can lead to concentrations of wealth and power incompatible with fairness and equal citizenship. For Rawls, unless the basic structure is properly regulated, it is illusory to expect that just distributions can sustain themselves over time.

4.5 Primary Goods

The two principles of justice define the **basic structure** of society, which in turn determines how **primary goods** are distributed among those who take part in social cooperation. These goods represent the various social conditions and resources that enable citizens "to develop and fully exercise their two moral powers and to pursue their determinate conceptions of the good" (*Justice as Fairness*, p. 57). They are, in other words, the things that individuals need in order to lead a complete life and to realize their personal life plans.

Rawls identifies five main categories of primary goods:

1. **Basic rights and liberties**, essential for the exercise of the two moral powers;
2. **Freedom of movement and occupational choice**;
3. **Powers and prerogatives** associated with positions of responsibility and public office;
4. **Income and wealth**, which provide the means to pursue a wide range of ends;
5. **The social bases of self-respect**, which sustain individuals' sense of worth and dignity.

Primary goods should be understood as **means** rather than ends. They are instruments that allow individuals to pursue their ultimate goals—goals that are determined by their **comprehensive doctrines**, whether secular or religious. The principles of justice thus specify how these means are to be distributed among different social classes, which, for simplicity, Rawls divides into the more advantaged and the less advantaged.

Rawls develops hypothetical "**index numbers**" representing the total amount of primary goods possessed by each social class. In the **original position**, citizens—placed behind the veil of ignorance—should choose, following the **difference principle**, those social arrangements (that is, sets of rules applied to the basic structure) which maximize the index value of the least advantaged class. In this way, inequalities are justified only insofar as they improve the condition of those at the bottom of the social hierarchy.

Rawls's approach, centered on the distribution of primary goods, has been notably criticized by **Amartya Sen**. Sen argues that Rawls's framework neglects the way in which individuals actually **use** or **benefit from** these goods. The well-being of individuals, Sen contends, ultimately depends on what they can *do* with the resources they have—their *achievements* and *real freedoms*. People differ greatly in their ability to convert the same bundle of goods into well-being: the elderly, the ill, or persons with disabilities, for instance, may derive far less benefit from a given set of resources than younger or healthier individuals. Sen therefore shifts the focus from the possession of goods to the development of **capabilities**—the real opportunities people have to do and be what they value.

Rawls responded that it is not correct to claim his theory ignores individual capacities. Citizens, he argues, acquire primary goods through their participation in social cooperation, which in turn depends on their **moral powers**—that is, their capacity to be equal and free participants in a fair system. However, this formulation implicitly excludes individuals who cannot fully take part in social cooperation—such as those with severe physical or mental disabilities. This exclusion, Rawls maintains, does not imply that society owes no solidarity to them; rather, he holds that the **central problem of political justice** concerns the fair terms of cooperation among citizens capable of participating as full members. The rights and needs of those who cannot do so must be addressed separately, through complementary principles of assistance and inclusion.

This clarification is crucial for understanding Rawls's use of the term "**least advantaged class**." In his framework, it does not refer to those at the very margins of society—the destitute, the excluded, or, in Christian language, "the last." Rather, it designates the group of **unskilled workers** who, while fully participating in the social and economic system, possess little or no human capital and depend primarily on their labor for income. Those at the margins, by contrast, require **special treatment**—ethical, social, and institutional—beyond the standard framework of distributive justice, to ensure that their basic dignity and inclusion are safeguarded.

4.6 The Difference Principle and the Selection of the Scheme of Cooperation

The **difference principle** is subordinate both to the principle of **equal basic liberties** and to the principle of **fair equality of opportunity**. It must therefore always coexist with these two. To understand how it operates, we can imagine a simplified situation in which society must choose among alternative **schemes of social cooperation**, each of which can be thought of as a set of rules defining the organization of productive activity, the division of labor, and the system of incentives that determines the distribution of rewards among social classes.

For illustrative purposes, let us imagine two representative groups: **workers** and **managers**. Each social scheme can be represented as a pair of values, *(a, b)*, where *a* denotes the index of primary goods assigned to the representative member of the working class, and *b* denotes the corresponding index for the managerial class. We shall assume that the society already satisfies the principle of fair equality of opportunity, so that the indices can be interpreted simply as the **per capita income** of the two groups.

Starting from a baseline scenario, we can suppose an egalitarian scheme where both groups enjoy the same income, say *(10, 10)*. However, other, more productive schemes of cooperation may exist that generate **greater benefits for both classes**, though with an unequal distribution of gains. For instance, a scheme might produce *(12, 18)* — higher incomes for both workers and managers, but with a proportionally greater advantage for the latter. According to the **difference principle**, this second scheme would be preferable to the egalitarian one, since it improves the position of the least advantaged (the workers) without violating the other principles of justice.

The difference principle thus **does not aim to eliminate inequalities**; rather, it allows them when and only when they serve to **improve the condition of the least well-off**. Inequalities are acceptable not as intrinsic goods, but as functional instruments for mutual benefit within a cooperative structure. This marks a sharp contrast with **utilitarianism**, which evaluates social arrangements according to the maximization of aggregate welfare or average utility, regardless of how benefits and burdens are distributed.

To see the difference, imagine yet another social scheme, *(11, 26)*, where the wealth of the managerial class increases substantially, but the workers' income falls compared to the previous case. Although this new scheme might maximize the **net welfare** of society, it would fail Rawls's test of justice because it improves the situation of the better-off **at the expense of the least advantaged**. The utilitarian logic of maximizing the sum or average of utilities allows such outcomes; the Rawlsian conception of justice as fairness does not.

What ultimately matters under the difference principle is that, among all possible schemes of cooperation, **society selects the one most advantageous for the least privileged group**. More efficient or dynamic social arrangements can be just, even if they produce greater inequalities, provided that the less advantaged genuinely benefit from the system as a whole. The difference principle can therefore be interpreted as a **principle of reciprocity** between social groups. It expresses the moral intuition that cooperation in a just society must be **mutually beneficial**—that citizens, despite differences in talent or social position, should regard their collaboration as fair because it is structured to ensure that everyone, and especially those worst off, gain from it.

4.7 Merit, and Talent

A society organized according to the **difference principle** does not reward individuals for their mere position in the social distribution of natural endowments, but rather for how they develop and employ these endowments in ways that contribute to the good of all. In Rawls's view, the basic structure of a just society should provide incentives for people to cultivate their capacities — through training, education, and effort — so that their talents become productive not only for themselves but for others as well. This reflects an implicit **principle of reciprocity** within the framework of justice as fairness: social cooperation should generate benefits that are shared, even if unequally, among all members of society.

Rawls famously describes the distribution of natural endowments — intelligence, creativity, physical ability, or any other innate trait — as a **common asset**. This does not mean that individuals lose ownership of their talents, or that the state acquires control over them. Rather, what is regarded as common is the **social value** of these endowments once they are organized within institutions of cooperation. Talents become complementary resources whose interaction produces a collective surplus. When structured by just institutions, the variety of natural abilities in society ceases to be a source of arbitrary privilege and instead becomes a foundation for mutual benefit.

Critics, however, have accused Rawls of advancing a theory of justice that is hostile to **merit**. Many hold that fairness requires individuals to receive what they "deserve" — that rewards should correspond to effort, ability, and contribution. The **meritocratic ideal** assumes that people are entitled to the fruits of their labor insofar as these stem from their personal achievements. From this perspective, the difference principle seems to undermine desert, since it limits how much advantage those with exceptional talents may rightfully enjoy for the sake of improving the condition of the least advantaged.

Rawls explicitly rejects the idea that natural talent itself can be a legitimate ground for moral desert. "No one," he observes, "deserves his place in the distribution of native endowments." A person does not merit being born with particular abilities any more than another deserves to be born without them. Talents are morally arbitrary facts of nature, not the result of personal virtue or effort. Hence, Rawls's conception of justice does not eliminate reward for effort; rather, it reframes it. What society rewards is not *the possession* of talent, but the **responsible and cooperative use** of talent — the willingness to refine one's abilities and apply them in ways that enhance the common good.

In this light, the Rawlsian society does not suppress ambition or excellence. Individuals are encouraged to pursue their self-realization, to aspire to positions of responsibility, and to excel in their chosen fields — industrial, political, scientific, or artistic. Those who exert greater effort and skill receive greater rewards, yet these rewards are constrained by the overarching demand that inequalities benefit all. The difference principle thus reconciles individual ambition with social justice: it channels personal success into a system of fair cooperation rather than unrestrained competition.

Libertarian thinkers such as **Robert Nozick** have sharply opposed this interpretation, arguing that treating natural talents as a "common asset" amount to a violation of **self-ownership** and of individual property rights. From a Nozickian standpoint, any redistribution that limits the rewards of the talented in order to assist others constitutes an unjust interference with personal liberty. Rawls, however, maintains that recognizing the social dimension of talent does not entail collective ownership of persons or their abilities. It only affirms that the benefits generated by natural advantages should be regulated by fair institutional rules — rules that respect equality of opportunity and ensure that the least advantaged also share in the gains of cooperation.

Ironically, in such a society, it is precisely the talented who stand to flourish most fully. When institutions guarantee fair opportunities and social recognition, individual excellence can be both rewarded and respected without degenerating into privilege. A Rawlsian social order would thus promote an environment in which talent is **discovered, cultivated, and socially valued**, where the state plays a legitimate role in fostering education, research, and creativity. Rather than suppressing merit, it transforms it — turning personal advantage into a **moral resource for collective advancement**.

Ultimately, Rawls's conception of merit is distinct from both **classical desert-based** theories and **utilitarian efficiency** models. It is merit within justice — a form of excellence that finds its legitimacy not in moral luck or natural hierarchy, but in its contribution to a shared system of cooperation that sustains the common good.

From a broader philosophical standpoint, Rawls's idea of the **distribution of talents as a common asset** anticipates later debates in **global justice**. Thinkers such as **Charles Beitz** and **Thomas Pogge** have extended this argument beyond national borders, suggesting that the natural and social resources of the world — including technological innovation, knowledge, and even the global climate — should be regarded as common assets of humanity. The Rawlsian principle of reciprocity among citizens thus becomes, in cosmopolitan form, a principle of solidarity among peoples. In this light, Rawls's moral architecture points toward a more encompassing vision of **justice for the global order**, where the accident of birth — whether of talents or of geography — no longer determines one's life prospects, and where the fruits of cooperation, at every level, are shared for the benefit of all.

4.8 The Choice of the Principles of Justice in the Original Position

Social contract theories, as developed by Hobbes, Locke, Rousseau, and Kant, have always aroused both fascination and skepticism. These theories sought to determine what kind of institutions or moral rules would emerge from the interaction of individuals in a hypothetical *state of nature*, that is, in the absence of any governing authority. The difficulty, as critics have often noted, lies in the very nature of this thought experiment: the state of nature is a philosophical fiction, and no actual person could plausibly consider themselves bound by principles derived from such an imagined situation.

Despite these criticisms, **John Rawls** takes the contractualist idea seriously. He reformulates it through his famous concept of the **original position**, which he conceives as a procedural device designed to make **Kant's categorical imperative** operational in the political realm. The thought experiment of the original position, for Rawls, allows us to identify objective and rational principles that can serve as the moral foundation of a just society. Some interpreters have gone further, suggesting that the original position can be understood as a kind of *archetype of justice* — a cognitive model that human beings unconsciously invoke when they attempt to resolve moral conflicts by "putting themselves in the place of others." Whether one takes it as a procedural device or as a psychological archetype, Rawls must still answer a crucial question: **why should rational individuals placed in the original position choose his two principles of justice rather than alternative principles, such as the utilitarian principle of maximizing average welfare?**

Rawls imagines the original position as a **deliberative procedure** in which hypothetical representatives of society choose the fundamental principles regulating its basic structure. These representatives are rational in the economic sense: they are capable of reasoning instrumentally about means and ends, they are not moved by envy, they display a strong aversion to risk and uncertainty, and they possess a sense of justice. Yet, they are motivated by **rational self-interest**. They are thus sophisticated agents, capable of evaluating alternative conceptions of justice according to rational criteria. Crucially, they deliberate **behind a veil of ignorance**: they do not know their class position, natural abilities, gender, race, or even their comprehensive conception of the good. Their deliberation focuses instead on **primary goods**, which are neutral with respect to different moral or religious doctrines and serve as general means for pursuing a variety of ends.

Given this epistemic situation, the parties are expected to choose principles that no one could reasonably reject once the veil is lifted. To be faithful to the expectations of the citizens they represent, they cannot bargain away certain non-negotiable conditions — above all, the **priority of basic liberties**. They would not, for example, trade off freedom of conscience or freedom of movement for economic gain. To do so would risk creating a society in which some individuals would find themselves unable to pursue their life plans at all. It is therefore unlikely that rational representatives in the original position would adopt principles that could render life intolerable for some members of society. For this reason, they would reject any arrangement that sacrifices basic liberties in exchange for greater material advantages.

The challenge arises in justifying the **difference principle**, according to which social and economic inequalities are acceptable only if they benefit the least advantaged members of society. Rawls must show that rational parties in the original position would adopt a **maximin strategy**—that is, a strategy that maximizes the position of the worst-off group. This requires that, when comparing possible social arrangements, they select the one in which the least advantaged achieve the highest possible level of primary goods, regardless of the relative condition of the more advantaged.

To see this more clearly, suppose the representatives must choose among the following possible social schemes:

A (10, 15), B (12, 18), and C (11, 26),

where the first number indicates the index of primary goods assigned to the least advantaged class, and the second refers to the most advantaged. When choosing between A and B, the parties would naturally prefer B, which benefits both groups. But the choice between B and C is more complex: scheme C offers far greater advantages for the rich, but slightly worse conditions for the poor. A purely rational agent—seeking to maximize expected outcomes—would want to know the probability of being rich or poor under each scheme, and would select the one with the **highest expected utility**.

Rawls, however, introduces three conditions under which it becomes rational to adopt the **maximin rule** rather than the expected-utility rule:
(a) the participants cannot assign meaningful probabilities to the possible social positions they might occupy (rich or poor);
(b) the level of primary goods guaranteed to the least advantaged under the optimal scheme is more than sufficient for a decent life; and
(c) the conditions of the least advantaged in schemes that promise much higher rewards to the rich would be so poor as to make life barely tolerable.

Under these assumptions, rational agents would prudently choose to maximize the minimum outcome—that is, to secure the best possible position for those who might end up at the bottom of society. This reasoning leads directly to Rawls's second principle of justice and to his rejection of utilitarianism, which allows severe inequalities if they raise average welfare.

Nevertheless, **critics such as John Harsanyi,** a Nobel prize in economics**,** have argued that Rawls's justification of the *difference principle* rests on conditions that are too restrictive and *ad hoc*. Harsanyi maintains that the parties in the original position, being rational and risk-averse but not infinitely so, would more plausibly apply the **principle of maximizing expected utility** rather than Rawls's maximin rule. In Harsanyi's formulation, if individuals are truly ignorant of their future position but know they are equally likely to occupy any role in society, the rational choice is to adopt the principle that maximizes the *average* well-being of all. This, in effect, yields a **utilitarian** outcome rather than Rawls's egalitarian one.

For Harsanyi, the key issue is not moral risk aversion but rational consistency: under uncertainty and symmetry of information, maximizing average utility is the only strategy that treats all persons as equals *ex ante*. Rawls's appeal to extreme risk aversion and to special constraints, Harsanyi argues, lacks a deeper justification and undermines the claim that his theory provides a general foundation for justice.

Yet this critique also reveals the distinctiveness of Rawls's project. By prioritizing *fairness* over aggregate welfare, Rawls explicitly rejects the reduction of justice to utility. His maximin rule embodies a moral commitment to protect the basic interests of each individual against the hazards of social inequality. Whereas Harsanyi seeks a rational consensus grounded in efficiency, Rawls aims for a moral consensus grounded in respect for persons as **free and equal moral agents**. The debate between Rawls and Harsanyi thus marks a fundamental tension in modern political philosophy—between **justice as fairness** and **justice as welfare**—and continues to shape contemporary discussions of equality, risk, and rational choice in theories of the just society.

The contrast between Rawls and Harsanyi, therefore, reveals not only a disagreement about rational choice under uncertainty, but also two profoundly different moral intuitions about the purpose of justice. For Harsanyi, justice is a matter of rational aggregation: when individuals deliberate under fair conditions, the principle that maximizes the average level of welfare for all is the most impartial and hence the most just. For Rawls, by contrast, the fairness of social arrangements cannot depend on the sum of satisfactions, but must rest on the moral equality of persons. What ultimately matters is not the efficiency of outcomes, but the assurance that each individual's basic rights and opportunities are protected regardless of their social position.

This tension between **utility and fairness**, or between **aggregate welfare and the protection of the least advantaged**, would later reappear at the heart of debates on **global justice**. Philosophers such as **Amartya Sen**, **Thomas Pogge**, and **Martha Nussbaum** sought to extend the Rawlsian framework beyond the boundaries of the nation-state, confronting the reality of global inequalities and the moral responsibilities of affluent societies toward the world's poor. Sen, in particular, challenged the Rawls–Harsanyi focus on ideal choice procedures, arguing that justice should not depend on hypothetical agreement but on the *comparative improvement of people's actual freedoms and capabilities*. Pogge, in turn, reinterpreted Rawls's principles in a cosmopolitan direction, claiming that the global order itself must be subject to principles of distributive justice to prevent structural harm.

Thus, the debate between Rawls and Harsanyi laid the conceptual foundations for a wider philosophical dialogue: whether justice should be grounded in rational efficiency or in moral equality, and whether its scope should remain domestic or extend globally. In this sense, the *original position* becomes not only a thought experiment about fairness within a single society, but a methodological precursor to contemporary reflections on **justice in the global order**—an order increasingly marked by interdependence, inequality, and the search for principles that can reconcile universal human dignity with the realities of global diversity.

4.9 The Law of Peoples

By *Law of Peoples*, John Rawls refers to a political conception of justice whose principles are meant to guide foreign policy, international law, and relations among nations. In his domestic theory of justice, Rawls had identified a political conception—*justice as fairness*—which applies to the basic structure of society and is endorsed by individuals who hold diverse conceptions of the good or different moral doctrines. In that domestic context, the relevant actors were individual citizens, and the goal was to enable them to pursue their own ends within a stable and just system of social cooperation.

In the international sphere, by contrast, the actors are nations, and the objective is to secure a just, stable, and peaceful global order. As will be seen, there are deep and important connections between the political conception of justice that governs domestic society and the one that applies to international relations.

Rawls's international project resonates with a Kantian intuition. In *Perpetual Peace*, Kant envisioned a federation of free republics governed by principles of right rather than by the logic of power. Rawls explicitly acknowledges this influence. Like Kant, he believes that a just international order is grounded in the moral autonomy of political communities, not in the imposition of a single global authority. However, whereas Kant's cosmopolitan ideal anticipated the eventual formation of a universal federation of states, Rawls remains cautious: he prefers a pluralist order of cooperative but independent societies, united by shared principles rather than by a supranational government.

It is important to note that Rawls deliberately speaks of a *Law of Peoples* rather than a *Law of States*. This terminological choice reflects his rejection of the state as a political entity possessing—even to a minimal degree—its own autonomous will or hidden agenda serving bureaucratic ambitions or the interests of a ruling elite. What Rawls envisions instead is a government that remains permanently under the control of its citizens, protecting their fundamental interests through responsible and transparent action. Consequently, governmental decisions should not be influenced by powerful economic interests or industrial lobbies.

For Rawls, the modern state, as it has historically developed, has always possessed sovereign powers—such as the right to declare war—that have often operated independently of the people's general will. Under the *Law of Peoples*, these sovereign powers, including those relating to war and certain aspects of internal autonomy, must be drastically curtailed.

The principal difference between *liberal peoples*—those who adhere to the *Law of Peoples*—and traditional states lies in their aims. Liberal peoples restrict their objectives according to a principle of reciprocity in their relations with other societies, whereas states have traditionally pursued their goals unilaterally—for instance, by attempting to convert others to their state religion or by expanding their territories to enhance economic power, prestige, or glory.

In developing his theory of international relations, Rawls classifies societies into five types:

1. **Liberal peoples**,
2. **Decent peoples**, which together constitute the so-called *well-ordered societies*;
3. **Outlaw states**, which operate outside the bounds of international agreements;
4. **Burdened societies**, meaning countries—often very poor—that lack the institutions, traditions, and human capital necessary to become well-ordered; and
5. **Benevolent absolutisms**, which respect human rights but deny their citizens any meaningful role in political decision-making.

The Social Contract among Liberal Peoples: Peace and Stability

To identify the political conception of justice that should underpin the *Law of Peoples*, Rawls again employs the device of the *original position*. Initially, the participants in this procedure are representatives of *liberal peoples*; *decent peoples* are incorporated only at a later stage. The condition for participating in the international contract is that a society must already meet certain minimum standards of internal justice. The underlying idea is that the international social contract must be preceded by a domestic one—established through the original position—that determines the principles of justice within each society.

In the international case, the actors—the representatives of peoples—deliberate under a *veil of ignorance* suitably adapted to this context: they do not know the size of their territory, their population, their natural resources, or their level of economic development. As in the domestic case, these representatives see themselves as free and equal moral agents.

It is also important to note that the participants are unaware of the comprehensive doctrines—religious, philosophical, or moral—professed by liberal peoples. The reason differs from that in the domestic case: a liberal society cannot itself be identified with any particular comprehensive doctrine. Such doctrines belong to individual citizens, associations, and communities within civil society. Accordingly, comprehensive moral or religious considerations have no place in the international contract.

According to Rawls, the delegates of liberal peoples would agree upon the following eight principles:

1. Peoples are free and independent, and their freedom and independence are to be respected by others.
2. Peoples are to observe treaties and undertakings.
3. Peoples are equal and are parties to the agreements that bind them.
4. Peoples are to observe a duty of non-intervention.
5. Peoples have a right to self-defense but no right to wage war except in self-defense.
6. Peoples are to honor human rights.
7. Peoples are to observe certain restrictions in the conduct of war.
8. Peoples have a duty to assist other peoples living under unfavorable conditions that prevent them from achieving a just or decent political and social regime (*The Law of Peoples*, p. 37).

These principles preserve the equality and independence of political communities while satisfying the criterion of reciprocity. Notably, there is no mention of redistributive principles like those found in the domestic social contract, apart from a general duty to assist societies in serious difficulty. The injustices within a particular society must be addressed internally. For Rawls, international redistribution would fail to correct the underlying mechanisms that structure a society's basic institutions. It is therefore more important to ensure respect for human rights, which are the essential precondition for any just society.

In addition to these eight principles, Rawls envisions three cooperative organizations: one to guarantee a fair system of international trade, one to serve as a world bank assisting countries in need, and one as a political body—a *Confederation of Peoples*. These correspond roughly to the GATT, the World Bank, and the United Nations. Together with the eight principles, they form the structural foundation of an ideal *Society of Peoples*.

A crucial feature of any principles of justice—domestic or international—is that they must ensure the stability of the society they govern. Citizens who participate in a system of social cooperation based on *justice as fairness* develop a strong sense of justice precisely because they grow up within such a system and see its principles consistently applied. This moral learning reinforces social cohesion and leads citizens not only to accept the principles of justice but also to act on them.

Rawls conjectures that a similar process operates among liberal peoples that endorse the eight principles of the *Law of Peoples*. Such peoples exhibit a moral character expressed through patriotism and pride in their history and traditions. They take pride in guaranteeing fundamental freedoms, sustaining a vibrant civil society, and securing general prosperity. They interact with others on the basis of mutual respect; their citizens are "fundamentally satisfied," their interests are compatible with those of other liberal peoples, and they have no desire to dominate, exploit, or impose their ideas on others. They recognize that their own security and well-being are best preserved in a general condition of peace. Accordingly, they are industrious, commercially minded, and oriented toward cooperation rather than conflict.

Rawls argues that a nation's predisposition toward or against war depends critically on its internal conditions. A country is immune to belligerence if its citizens enjoy extensive basic liberties, if social and economic inequalities remain moderate, and if there exists broad public debate on major political and economic issues—that is, if genuine democratic practice prevails. Conversely, states prone to conflict are those where sharp divisions separate the people's interests from those of the ruling classes, where governments serve powerful economic lobbies, or where political life is driven by rigid ideological motives.

Rawls thus concludes that peace is possible among modern liberal democracies, even though they may at times engage—often for questionable reasons—in conflicts with non-democratic states or in covert operations to subvert foreign governments. Such conflicts, he contends, stem not from democratic ideals but from oligarchic tendencies, entrenched interests, and internal injustices that distort liberal societies.

For lasting peace among democracies, Rawls identifies several internal conditions:
a) Equal opportunities, especially in education, without which not all groups can participate effectively in public deliberation. For Rawls, democracy is above all the exercise of public reason.

b) A fair distribution of income and wealth, since economic concentration leads to political domination.
c) Policies that guarantee stable and lasting employment, as insecurity undermines self-respect and civic belonging.
d) Social and health assistance ensuring basic welfare for all.
e) Public financing of political parties, to ensure independence from powerful economic interests.

Rawls's conjecture is that constitutional liberal democracies, devoted to the practice of public reason and supported by a basic structure that secures social justice, will not wage war against one another.

Historical evidence appears to confirm this. Since the nineteenth century, liberal democracies—though often at war with non-democratic states—have never fought each other. The great wars of the nineteenth century, such as the Napoleonic and Bismarckian wars and the American Civil War, were not conflicts among democracies. In both world wars, democratic nations were allies. Later conflicts involving the United States, such as Korea and Vietnam, were not wars between democratic states.

Possible exceptions include covert operations by the United States that toppled democratic governments, such as those of Allende in Chile, Arbenz in Guatemala, or the Sandinistas in Nicaragua. Rawls attributes such actions to the influence of powerful lobbies shaping foreign policy. By contrast, the 1991 Gulf War—waged by a coalition of thirty-four states under UN authority against Iraq—targeted a non-democratic regime. The wars in Afghanistan (2001) and Iraq (2003) likewise involved interventions against non-democratic states.

The Extension of the Social Contract to Decent Peoples

To extend the social contract to non-liberal peoples—and thus include them within the *Society of Peoples* governed by the eight principles—Rawls asks how far liberal peoples should go in tolerating non-liberal societies. *Toleration* here means refraining from coercive sanctions and accepting such peoples as full members of the Society of Peoples, with corresponding rights and duties, including the obligation to justify their political actions through public reason.

Two main criteria must be met for a non-liberal society to qualify as a full member.

First, the society must be non-aggressive and recognize that its legitimacy depends on diplomacy, trade, and other peaceful means. Even if a dominant religion influences the political structure—for example, by reserving key offices for religious leaders—it must nonetheless respect the political and social independence of other societies and refrain from coercive proselytism.

Second, the society must satisfy three internal conditions:

1. It must guarantee its members a core set of human rights, understood as a subset of those typical of democratic societies—namely,

    - the right to life, ensuring a minimum level of subsistence;
    - freedom from slavery and oppression;
    - freedom of conscience;
    - formal equality (treating like cases alike); and
    - the right to personal property.
    These rights, Rawls emphasizes, are not peculiar to the West but exist at least as legitimate aspirations in all human cultures.
2. It must have a legal system that imposes duties and obligations which citizens accept because they recognize them as part of a shared conception of justice, not merely because they are enforced.
3. Judges and officials must sincerely believe that the legal system is guided by a genuine and publicly accepted conception of justice.

When these conditions are met, the society may join the *Society of Peoples*. Rawls calls such societies *decent hierarchical societies*. Although they lack full democracy, they include mechanisms of hierarchical consultation whereby the views of various social groups are heard before major political decisions are made. They therefore differ from liberal societies yet retain a morally respectable status.

Critics argue that the human rights listed above can in practice be protected only within democratic systems. Rawls replies that decent hierarchical societies already exist—for instance, in parts of the Muslim world—and that human rights are not an invention of Western civilization but express aspirations inherent in human nature. He maintains that such rights can, and eventually will, prevail even in societies that currently deny them. There are several Muslim-majority countries that, while not perfect, have made substantial progress toward building societies that respect human rights, rule of law, and pluralism. Malaysia, Indonesia, Jordan, Morocco, to name some, stands out.

Both liberal and decent peoples—the well-ordered members of the *Society of Peoples*—share the long-term goal of encouraging *outlaw states* to reform and come to respect the *Law of Peoples*. How this transformation occurs depends on circumstances and may take different forms, including moral pressure, public denunciation, or economic sanctions. Although authoritarian rulers may dismiss such measures as Western propaganda, Rawls insists that they cannot easily evade the issue, since human rights possess a universal character rooted in the moral nature of humanity itself.

The Duty Toward Burdened Societies

Well-ordered peoples have a duty to assist those societies that, while not aggressive, are unable—due to historical traditions, lack of technical knowledge, or insufficient institutional development—to become well-ordered on their own. The level of well-being among such societies may vary greatly: some are extremely poor, others less so, but what unites them is the absence of a shared political conception of justice.

According to Rawls, the problem of international assistance is not one of **equalizing per capita income or wealth** among citizens of different countries. Rather, it is one of **helping societies establish just institutions** that will enable them to join the *Society of Peoples* and thus contribute to a stable and peaceful international order.

It should be noted, incidentally, that Rawls recognizes that even well-ordered societies can differ widely in their levels of income and wealth per capita. In such societies, which have already achieved just political and social institutions, income and wealth depend primarily on three factors: the rate of capital accumulation, the degree of technological progress, and demographic control. Rawls, following John Stuart Mill, maintains that saving and capital accumulation may slow down or even cease in societies that have reached a high level of well-being, where the capitalist spirit weakens and gives way to the pursuit of moral and cultural ends. However, there is no evidence today that the wealthiest countries are moving toward a stationary state; and even where savings are declining, this is not because societies have achieved a sense of satiety or reconsidered the capitalist model of development.

Returning to the main theme, Rawls argues that it is **not necessary** for a burdened society to reach a particular level of income or wealth before it can establish just political and social institutions. Economic conditions can vary from society to society depending on the preferences and attitudes of their citizens, and are therefore, to some extent, **independent of the institutions that regulate justice**.

Rawls thus believes that the proper way for well-ordered peoples to assist those in difficulty is **not through simple transfers of wealth**, since financial aid alone will not correct the unjust institutions that characterize these societies. It is far more effective, he argues, to insist on the **respect for human rights** and to promote institutional reform, rather than to focus on redistributive policies from rich to poor countries.

Rawls endorses the important findings of **Amartya Sen's work** on the relationship between freedom, human rights, and economic development in poor countries. Contrary to earlier assumptions, Sen demonstrated that poor societies handle crises and famines much more effectively when they possess decent institutions and respect for human rights. Such rights should therefore not be regarded as a luxury reserved for affluent nations but as essential conditions for development itself.

The respect for human rights can also play a crucial role in easing **demographic pressures** within a society. A decisive aspect here is the improvement of women's social status. For instance, China has attempted to address overpopulation through invasive social policies restricting family size—policies that deeply infringe on individual freedom. In contrast, the Indian state of **Kerala** adopted an alternative approach: by expanding women's rights, improving their education, and increasing their participation in political and economic life, it achieved a significant reduction in birth rates without coercion.

In offering assistance to burdened societies, Rawls insists that **well-ordered peoples must not act paternalistically**. Instead, they should help these societies find, within the bounds of their own traditions, ways to reform their economic and social institutions. Once certain institutional goals have been achieved, such societies must be allowed to progress autonomously, even if their income levels remain modest. In defining the objectives of assistance, priority should always be given to the **basic structure of society**, rather than to per capita income or economic growth as such.

Rawls, Pogge, and Beitz: The Scope of Global Justice

Rawls's conception of the *duty of assistance* occupies a cautious and intermediate position in the contemporary debate on global justice. His vision is profoundly **political rather than economic**, aiming to secure a fair framework among self-governing peoples rather than to eliminate global inequalities in wealth. This limited humanitarian duty is rooted in his commitment to respecting national self-determination and cultural pluralism: each society must, ultimately, be the author of its own path toward justice.

By contrast, **Charles Beitz** and **Thomas Pogge**, both deeply influenced by Rawls, challenge this restriction. Beitz, in his *Political Theory and International Relations* (1979), argues that the principles of distributive justice should extend beyond national borders, because the world economy constitutes a **single cooperative system** in which the fortunes of nations are interdependent. For Beitz, Rawls's framework unjustifiably stops at the borders of the state: if justice as fairness governs cooperation within societies, it should also govern the global structure that binds societies together.

Thomas Pogge, in *World Poverty and Human Rights* (2002), develops this argument further. He maintains that the existing global economic order is not a neutral background but a **causal structure** that perpetuates poverty and deprivation. The affluent societies, by benefiting from unfair trade, resource, and financial arrangements, share **negative responsibility** for the suffering of the poor. In this view, the duty of assistance cannot be reduced to charity or institutional guidance; it must entail **rectifying the injustices** of the global system itself.

Rawls, however, resists such cosmopolitan extensions. For him, justice is **context-bound**: its principles apply to distinct societies that have freely chosen to cooperate within the limits of reciprocity. To universalize distributive justice would, in his view, undermine political autonomy and the moral pluralism that defines the *Law of Peoples*. Rawls's global vision therefore mirrors Kant's federation of free and equal states—anchored in mutual respect, human rights, and peaceful coexistence, but without a world government or a global distributive mechanism.

In this respect, Rawls's *duty of assistance* articulates a **moral minimalism** within a broader architecture of international justice: it seeks to secure the conditions of political decency and self-sustainability for every people, not material equality across the globe. Yet this minimalism is not moral indifference. It represents, instead, a form of **political cosmopolitanism**—a belief that the world can be ordered justly through institutions that respect both difference and moral reciprocity.

From Kant's cosmopolitan right to Rawls's *Law of Peoples*, the evolution of modern political thought reveals an ongoing effort to reconcile the moral unity of humanity with the political plurality of the world. While Beitz and Pogge push the boundaries toward a more demanding cosmopolitan egalitarianism, Rawls's approach remains anchored in the **moral discipline of political legitimacy**: justice arises not from an abstract ideal of world citizenship, but from the fair cooperation of free and responsible peoples.

This balance between moral universalism and political autonomy defines the core of what we may call a *Philosophy for Global Justice*—a vision of a moral world order that is neither utopian nor resigned, but guided by the conviction that institutions, when grounded in fairness and human dignity, can progressively civilize the global condition.

This discussion of Rawls's *duty of assistance* closes the arc linking domestic justice to the global order. It also prepares the ground for contemporary debates examined in the next part of this book—debates where the moral insights of Rawls, Beitz, and Pogge encounter the urgent realities of **human rights, migration, climate change, and global sustainability**. The question that follows is whether the moral world order envisioned by Rawls can endure in a world increasingly shaped by interdependence, inequality, and ecological limits

4.10 Justice Between Generations: The Just Savings Principle

One of the most subtle and problematic dimensions of Rawls's theory concerns **justice between generations**, which he addresses through what he calls the **principle of just savings**. In Rawls's view, each generation has a duty to accumulate and preserve sufficient resources to sustain a just social order over time. Savings thus perform a crucial function: they are not merely instruments of economic growth, but the moral means of maintaining the conditions of justice for future citizens. Once a just institutional framework and adequate capital base are secured, the need for further accumulation gradually disappears.

By *capital*, Rawls means not only physical goods—factories, machines, or infrastructure—but also **knowledge, culture, and social institutions**. These form part of the moral inheritance that one generation passes to the next. In this sense, savings are as much cultural and institutional as they are economic.

As Rawls famously puts it, *"the principle of just savings holds between generations, while the difference principle holds within generations."* This distinction reflects the two temporal dimensions of justice: the intragenerational distribution of benefits among contemporaries, and the intergenerational transmission of the conditions that make justice possible at all.

Two Models of the Original Position

Rawls distinguishes between two possible interpretations of the **original position** when extended to intergenerational justice:

1. **Model (a):** Representatives of all generations deliberate together, under a veil of ignorance, without knowing to which generation they belong.

2. **Model (b):** Only representatives of one generation deliberate; they know they are contemporaries, but do not know which generation they belong to in historical time.

Rawls rejects **Model (a)** because it presupposes the existence and size of all generations as fixed and known, which is both empirically and conceptually implausible. Such a view would ignore the relationship between **demographic policies and saving behavior**—a crucial factor in real societies. Instead, he adopts **Model (b)**, limiting the deliberation to contemporaries.

However, this model generates a serious problem. In the original position, the parties are motivated by **rational self-interest**. Since they will not personally benefit from sacrifices made for future generations, they have no reason to adopt a saving principle. In other words, rational agents would consume rather than save, leaving no provision for their descendants.

To resolve this, Rawls modifies his own model. He assumes that the parties represent **family lines**, linked by "ties of sentiment" across at least two generations. These intergenerational bonds are intended to introduce a form of **limited altruism**, sufficient to motivate savings for the immediate future. Yet this move, while ingenious, is conceptually troubling.

The Limits of Altruism and the Problem of Motivation

The introduction of altruistic feelings within the original position **contradicts the assumption of self-interest** on which the entire contractual framework is built. Altruistic motivation, by its nature, does not extend indefinitely into the future—it cannot encompass **remote generations**. As a result, Rawls's model largely neglects long-term processes such as **environmental degradation, climate change, and the depletion of natural resources**, whose effects extend far beyond the horizon of direct familial concern.

Moreover, because only contemporaries are represented in the chosen model, the framework risks collapsing into a theory of **duties of the present generation toward the next**, rather than genuine **justice between generations**. Even if generation *k* agrees to save at a certain rate, there is no guarantee that later generations will do likewise. In a sense, Rawls's model reproduces an *n*-generation **prisoner's dilemma**: each generation would be willing to save only if it could trust that previous generations had done so and that future ones would reciprocate. Yet such coordination is impossible within the single-generation deliberation of Model (b).

The difficulty is structural. The postulation of altruism within a model built on self-interest **undermines the coherence of Rawls's contractarian logic**. The *veil of ignorance* was designed precisely to ensure impartiality among self-interested agents. To rely on moral sentiment to sustain intergenerational justice introduces an inconsistency between **Humean motivational premises** and **Kantian moral impartiality**.

Rawls's Revised Proposal

In response to these objections, Rawls proposes a revised interpretation of the just savings principle. He suggests that the parties in the original position would adopt **the saving rate they would wish previous generations to have followed**, and that they would also want future generations to continue following.

This revision introduces a **temporal reciprocity**: justice between generations is grounded in the idea of fairness across time, rather than altruistic sentiment. The principle thus becomes an expression of **universalizability**, akin to Kant's categorical imperative—each generation is to act by principles it would be willing to have others, past and future, adopt.

Under this formulation, the just savings principle is **mutually beneficial** for all generations except the first, which bears the initial cost of building just institutions. Once a stable and just basic structure is achieved, the burden of savings can decrease, and future generations inherit not only material wealth but also a framework of justice.

Still, even this elegant solution does not entirely overcome the motivational gap. The representatives of one generation, even behind the veil of ignorance, may remain indifferent to the fate of distant future people. The contractual model, oriented toward rational agreement among contemporaries, has difficulty extending moral concern to those who cannot reciprocate—a problem that continues to haunt theories of intergenerational and ecological justice.

Parfit's Challenge: The Non-Identity Problem

Philosopher **Derek Parfit** sharpened this difficulty through what he called the **Non-Identity Problem**. Parfit observed that many of our actions—such as environmental policies or decisions about resource use—determine not only the quality of life of future people but **which people will exist at all**. This implies that we cannot easily claim to harm or wrong future generations, since our actions also determine their very identity.

Parfit's argument reveals a deep **metaphysical tension** in Rawls's model. The original position assumes that the same individuals can evaluate justice across generations, yet in reality, future persons are not yet identifiable subjects of harm or benefit. Parfit therefore challenges the very notion of contractual reciprocity between temporally distinct populations. For him, what matters is not the identity of future individuals, but the **impersonal value** of ensuring that future lives are worth living.

This shift—from justice between specific persons to the morality of outcomes—moves beyond Rawls's contractarian limits. Parfit's view demands that we care about the long-term future *as such*, independently of personal ties or hypothetical agreements. It is an ethics of **responsibility without reciprocity**, one that extends moral consideration to the unborn and to the integrity of the planet itself.

Toward an Intergenerational Global Justice

The debate between Rawls and Parfit illuminates a broader philosophical challenge: how can justice remain both **politically grounded** and **morally inclusive** in a temporally extended world? Rawls's *Law of Peoples* provides a framework for justice among nations, but his principle of just savings struggles to secure justice **among generations**. Parfit's critique suggests that we must think of global justice not merely as fairness in space but also as **fairness through time**.

In this expanded moral horizon, the duty of justice includes responsibility for the **conditions of human survival**—for the sustainability of the natural and institutional order that future generations will inherit. Here the concept of a *Philosophy for Global Justice* acquires its full meaning: it is not only about the relations among peoples and states but also about the **continuity of humanity itself**, about preserving the moral and ecological foundations of a just world for those who will come after us.

This discussion of intergenerational justice points directly to the ethical foundations of **sustainability and global responsibility**. The tension between Rawls's institutional minimalism and Parfit's impersonal ethics anticipates the contemporary search for a normative framework capable of addressing climate change, environmental degradation, and intergenerational inequality.

Appendix A: The Social Contract in Hobbes, Locke, and Rousseau

The idea of the social contract has played a central role in shaping modern political philosophy and the development of democratic thought. Through it, thinkers such as Thomas Hobbes, John Locke, and Jean-Jacques Rousseau sought to explain the origin of political authority, the nature of human freedom, and the foundations of legitimate government. Though united by a common concern for order and justice, their theories reflect distinct conceptions of human nature and society.

Thomas Hobbes

Thomas Hobbes (1588–1679) believed that human beings are driven by instinct and appetite, much like other animals. He viewed human behavior as determined by natural causes and motivated by self-interest and the instinct for survival. Although people are rational, they are also distrustful, competitive, and inclined to conflict. Hobbes described this natural condition as one where each person is against the other—*homo homini lupus* ("man is a wolf to man").

Life in the state of nature, according to Hobbes, is "solitary, poor, nasty, brutish, and short." It is a pre-moral state without law or justice, where everyone has a right to everything and nothing is secure. Because human beings naturally desire peace and order, their fear of violent death leads them to use reason to find a way out of this condition.

To escape the insecurity of the state of nature, individuals agree to surrender all their natural rights and freedoms to a common authority. This agreement, called the social contract, creates a sovereign power that guarantees protection and stability. Before the contract, people were morally free to do whatever they wished; but once they make a covenant, breaking it becomes unjust. Justice, for Hobbes, is not natural but artificial—it arises from human agreements within civil society.

The result of this social contract is the establishment of an absolute authority whose duty is to protect life and property. Hobbes supported absolutism, arguing that "Covenants, without the sword, are but words." The sovereign's power must be unlimited, because only absolute authority can maintain peace and prevent a return to chaos. In Hobbes's theory, individualism, materialism, utilitarianism, and absolutism are closely connected.

John Locke

John Locke (1632–1704) offered a very different version of the social contract. He agreed that people are naturally free and equal, but he did not see the state of nature as violent or chaotic. Instead, Locke described it as a peaceful and reasonable condition—a "Golden Age"—where natural law governs and people respect each other's rights. However, property in the state of nature was insecure because there was no common law, no impartial judge, and no authority to enforce justice.

Locke placed great importance on property. He believed that private property is created when an individual mixes labor with natural resources. Because the protection of property was uncertain in the state of nature, people agreed to form a political society. Through the social contract, they transferred only some of their rights—those necessary to preserve order and enforce the law—while retaining their natural and inalienable rights to life, liberty, and property.

By entering civil society, individuals gained what they lacked before: known laws, impartial judges, and an executive power to enforce them. The purpose of government, Locke argued, is to protect natural rights. When a government fails to do so, its laws lose validity and the people have the right to replace it. Unlimited sovereignty is therefore contrary to natural law. Locke supported a constitutional and limited government based on the consent of the governed and directed toward the common good.

Jean-Jacques Rosseau

Jean-Jacques Rousseau (1712–1778) further transformed the theory of the social contract. He rejected both Hobbes's view of humans as selfish and the Christian idea of original sin. Rousseau believed that people are naturally good, compassionate, and capable of justice, but that social institutions can corrupt these qualities. Human nature, in his view, is shaped by the social and political environment.

In the state of nature, people lived simply, freely, and equally. The development of private property, however, introduced inequality and dependence, leading to moral and social decay. The social contract, for Rousseau, is meant to restore freedom and equality on a moral foundation. Individuals agree to join together in a community where each person gives up some natural freedom in exchange for collective security and participation in forming the general will.

Unlike Hobbes, Rousseau believed that sovereignty must always remain with the people. The only legitimate government is one that expresses the *general will*—the collective will of the people directed toward the common good. The general will is not the same as the sum of individual interests (the "will of all"); rather, it represents what is best for society as a whole. It is expressed through laws that apply equally to all citizens and aim at the common good.

When individuals obey such laws, they are not submitting to an external authority but to a rule they have created together. Rousseau famously wrote that individuals who resist the general will may be "forced to be free," meaning that they can be compelled to follow laws that express their own collective freedom. In Rousseau's view, true freedom is not the ability to do whatever one wants, but the ability to obey laws one has helped to make. The social contract thus creates a political community based on equality, shared participation, and moral autonomy. Through it, individuals achieve both freedom and unity within the framework of the general will.

Although Hobbes, Locke, and Rousseau all based their political theories on the idea of a social contract, they differed sharply in their understanding of human nature, the state of nature, and the purpose of government. For Hobbes, human beings are naturally selfish and fearful, requiring an absolute authority to maintain peace and order. Locke, by contrast, saw humans as rational and cooperative, governed by natural law and endowed with inalienable rights to life, liberty, and property; government exists only to protect these rights and remains limited by the consent of the governed. Rousseau, finally, viewed humans as naturally good but corrupted by society. His social contract sought not only security but moral and political freedom through participation in the general will. In short, Hobbes's contract creates authority, Locke’s protects rights, and Rousseau’s establishes freedom and equality within a moral community.

Appendix B: Immanuel Kant and the Project for Perpetual Peace

Immanuel Kant (1724–1804) wrote the *Project for a Perpetual Peace* (*Zum ewigen Frieden*) in 1795. That same spring, the Peace of Basel was signed, by which Spain, Holland, and Prussia formally recognized revolutionary France as a major European power. The political upheavals of the time deeply shaped Kant's reflection on war, peace, and the moral destiny of humankind.

Like Hobbes, Kant maintains that the State arises from a pact among individuals. Human beings, in his view, agree to coexist under common laws because they find such an arrangement both rational and advantageous. The State is thus a moral and legal order born from the decision to transform natural conflict into peaceful cooperation.

In the *Project*, Kant extends this social contract logic to the international level. Just as individuals in the state of nature reach an agreement and create the State to secure peace, so States—conceived as "large individuals"—should agree among themselves to form a federation that guarantees perpetual peace.

From the Psychology of War to Its Social Causes

Before Kant, many thinkers had addressed the problem of war—from Erasmus of Rotterdam (1466/69–1536) to the Abbé de Saint-Pierre (1658–1743). Their proposals, however, largely focused on the moral failings of rulers: ambition, aggression, or vanity. They culminated in appeals to the good will of princes. Kant's diagnosis is radically different. Wars, he argues, do not stem from the bad intentions of sovereigns but from the **social structure of absolutism itself**—the political order of the *Ancien Régime*, which perpetuates inequality and conflict.

In the *Critique of Judgment*, Kant articulates the principle behind this argument:

"Only under one condition can nature achieve its final intent: and it is when, in a totality called civil society, a legal power is formed, which is opposed to the impairment of the freedoms of individuals in their mutual contrast; since the maximum development of natural dispositions can only take place in such a society."

Human beings, Kant observes, are both sociable and unsociable. Left to their instincts, they tend toward conflict—each being "a wolf to another man." Yet precisely through this tension, when forced by necessity into lawful coexistence, humanity progresses. The State, as a structure that limits selfishness and channels freedom into law, becomes the instrument through which human potential is realized. Institutions are what matter.

Peace as a Human Construction

"The state of peace among men, who live side by side, is not a natural state," writes Kant. Peace, therefore, must be constructed: it is an *artificial achievement*, a conscious moral and political effort. By nature, both individuals and States exist in a condition of belligerence. To move beyond this, Kant identifies three fundamental conditions for achieving perpetual peace:

1. Every State must have a republican constitution.

2. A federation of free States must be established.
3. A universal right of hospitality must be recognized.

The Republican Constitution

For Kant, perpetual peace is inconceivable unless States are republican. Yet republicanism cannot be imposed from outside—it must arise from within, through the internal moral and political development of peoples. This makes the road to peace long and uncertain.

A republic, in Kant's sense, is characterized by **freedom**, **equality**, and **representation**. Freedom here does not mean unrestrained liberty, as in the state of nature, but the freedom of rationally accepted laws. Equality follows from this, since laws that are rational apply equally to all citizens. Finally, republics are representative governments: power is mediated through elected officials who express the will of the people. When direct rule bypasses this mediation, despotism re-emerges.

Why does republicanism promote peace? Kant's reasoning is strikingly modern:

"To decide whether there should be war or not, the consent of the citizens is required. Having to decide themselves to suffer all the calamities of war—the fighting in person, the financial burdens, the ruins and debts—they will think long before beginning such an evil game. In contrast, where the ruler is not a fellow citizen but the owner of the State, he may declare war for trivial reasons, since it will not disturb his banquets or his courtly pleasures."

Thus, the republican form of government is both the moral and institutional foundation of perpetual peace.

A Federation of Free States

The second requirement is the creation of a federation among States. Here, Kant's **contractual vision** of politics becomes clear. States, like individuals in the state of nature, can commit injustice simply by existing side by side without a higher law to regulate their relations.

However, a "universal republic"—a single global State—is not conceivable, because it would destroy the very sovereignty that defines political identity. Therefore, the only feasible path is a **voluntary federation of free and equal States**, united not under coercive authority but by mutual agreement.

This raises a paradox: how can States leave the state of nature without surrendering sovereignty? Kant's answer reveals his Enlightenment optimism. Even when nations wage war, they appeal to legality. Aggressors seek justification in legal terms; they feel compelled to present a *rightful cause*. From this, Kant infers that the **idea of law itself** will gradually prevail—imposing limits on violence and guiding humanity toward a lawful condition among nations.

The Right of Hospitality

The third condition for perpetual peace is the **right of hospitality** (*ius cosmopoliticum*). Kant's reasoning anticipates the moral foundations of cosmopolitanism. All human beings share the Earth; the planet belongs to everyone. Birthplace is accidental, and therefore all people have a right to move freely and to be received peacefully elsewhere.

This right, Kant clarifies, is not a right to permanent residence but a *right of temporary visitation*:

"It is the right to join the company of others under the right of common possession of the surface of the Earth, on which, being spherical, men cannot disperse indefinitely but must finally endure to stand side by side."

Kant contrasts this ideal with Europe's colonial injustices. The so-called "civilized" powers—especially commercial nations—have violated the right of hospitality through conquest, enslavement, and plunder. "America, the Negro countries, the Spice Islands, the Cape of Good Hope," he writes, "were treated as lands belonging to no one, their inhabitants ignored."

The nobility of the ideal is undeniable, yet Kant recognizes its fragility: republics do not yet exist everywhere; States are reluctant to cede sovereignty; and the right of hospitality remains widely violated. To strengthen his case, Kant introduces a final argument—the **guarantee of perpetual peace**.

Nature as the Guarantor of Peace

The guarantee of peace, paradoxically, comes not from human will but from nature itself:

"What provides this guarantee is nothing less than nature, the great creator (natura daedala rerum), from whose mechanical course we see the finality shine, which from discord among men produces concord even against their will."

This "destiny" or "providence" expresses Kant's teleological faith in history. Despite wars and antagonisms, human progress continues. Even destructive figures such as Napoleon, he observes, unintentionally contribute to civilization: the Napoleonic wars, though ruinous, brought about the abolition of feudalism and the spread of civil law.

Politics and Morality Reconciled

For Kant, peace is ultimately a **moral ideal**. To achieve perpetual peace, politics must conform to morality. Against Machiavelli, who separated politics from ethics, Kant envisions their reconciliation. The true statesman must act as a moral man; the pursuit of peace must be guided not by expediency but by duty.

Kant's *Project for a Perpetual Peace* thus represents one of the earliest and most profound attempts to formulate a **moral philosophy of world order**. His vision of a federation of free republics bound by law, mutual respect, and hospitality remains one of the foundational inspirations for modern theories of **global justice** and **international cooperation**—anticipating, in many ways, the spirit of Rawls's *Law of Peoples* and the continuing philosophical search for a just and peaceful global community.

Appendix C: Affirmative Action and the Challenge of Non-Ideal Justice

John Rawls devoted most of his philosophical work to what he called **ideal theory**—the study of what a *perfectly just society* would look like under conditions of full compliance with the principles of justice. Ideal theory, however, leaves open an essential question: *What should we do when we live in an unjust society?* The answer to that question belongs to what Rawls termed **non-ideal theory**—the practical ethics of justice in imperfect circumstances.

Affirmative action arises precisely within this **non-ideal domain**. It is a policy designed to address the **unjust consequences of an unjust history**. The question of whether affirmative action is itself just, or merely a new form of unfairness, lies at the heart of moral and political reflection in societies like ours—societies that have formally abolished legal discrimination but still bear the deep structural scars of past injustice.

Historically, affirmative action became a major public issue in the early 1970s, following the final elimination of legal segregation in the United States. The landmark case of *Regents of the University of California v. Bakke* (1978) made the question of preferential hiring and admissions a constitutional matter. The case arose when **Allan Bakke**, a white applicant, was twice denied admission to the University of California, Davis Medical School, which reserved sixteen of one hundred seats for minority applicants under a special admissions program. Bakke claimed that this policy constituted **reverse discrimination**, violating the **Equal Protection Clause of the Fourteenth Amendment** and **Title VI of the Civil Rights Act of 1964**.

The Supreme Court's decision was deeply divided and produced a complex outcome. The Court held that the **use of rigid racial quotas** in admissions was unconstitutional and that Bakke should be admitted. However, a majority also affirmed that **race could be considered as one of several factors** in admissions decisions to promote diversity and to remedy the effects of past discrimination.

Justice **Lewis Powell's** opinion became the controlling one: he argued that while racial quotas violated equal protection, **diversity in education was a compelling interest** that justified limited use of race as a criterion.

Although **Rawls never published a specific treatment of affirmative action**, it is known that he privately expressed his view that its constitutional defense was morally important. His silence on the topic may reflect his methodological commitment to ideal theory, which leaves little room for the kinds of corrective policies that non-ideal conditions demand.

Nonetheless, Rawls's framework offers a powerful set of tools for thinking about affirmative action. Segregation and racial discrimination clearly violate his **first principle of justice**, which guarantees *equal basic liberties for all citizens*. Yet the debate over affirmative action begins precisely when such explicit forms of discrimination have been abolished. What remains is a deeper, structural inequality—one that persists even when the law is formally neutral.

Rawls's Framework: Fair Equality of Opportunity

Within Rawls's conception, the injustice that affirmative action seeks to address is a particular form of the **failure of fair equality of opportunity**. In *A Theory of Justice*, Rawls insists that equal opportunity requires more than the mere absence of legal barriers; it demands that social and economic inequalities be arranged so that individuals with similar talents and motivation have genuinely equal prospects of success, regardless of their social background.

In societies marked by historical and structural discrimination, this condition is far from being met. Those born into poverty cannot compete effectively with the well-off for desirable positions; they are less likely to receive quality education, social support, or even the motivation to pursue success. Formal equality of rights coexists with **substantive inequality of prospects**.

For African Americans, this injustice is compounded by centuries of slavery, followed by a hundred years of legally enforced segregation, and by the lingering psychological and social prejudices that mark both minority and majority populations. The result is a **systemic disadvantage** that reproduces itself across generations—a social inertia that Rawls's ideal model cannot correct by itself.

From a Rawlsian perspective, affirmative action can thus be understood as a **temporary, corrective mechanism**designed to restore the conditions for genuine equality of opportunity. It seeks to break the self-reinforcing cycle of deprivation that keeps marginalized groups at a structural disadvantage, even in the absence of overt discrimination.

Thomson's Justification: Preferential Hiring as Moral Repair

In her influential essay *"Preferential Hiring"* (1973), **Judith Jarvis Thomson** provides a rigorous moral justification for affirmative action, one that complements Rawls's normative framework. Thomson distinguishes between **compensation** for past wrongs and **rectification** of present inequality. Preferential hiring, she argues, is not a form of collective guilt or retribution but an act of **institutional moral repair**—a means of counteracting the enduring effects of injustice that continue to shape the structure of opportunity.

Thomson's defense is grounded in a principle of fairness: if the current system continues to advantage one group due to historical injustice, fairness requires intentional measures to correct that imbalance. In this sense, preferential policies are not reverse discrimination but **structural justice** in action—they aim to reestablish the moral baseline that an unjust history has destroyed.

Her argument also addresses the **psychological dimension** of oppression. In societies long divided along racial lines, social identity shapes aspiration, self-confidence, and perceived legitimacy. Disadvantage is not merely economic or legal; it is also **internalized**. Preferential hiring and admissions policies, by opening access to spaces of privilege and authority, serve not only distributive but also **symbolic and restorative functions**. They help reshape the social imagination of equality.

Rawls and Thomson: Justice as Correction

Rawls's principles, when interpreted through Thomson's lens, can be extended into the domain of non-ideal theory. The **difference principle** already permits inequalities only when they work to the advantage of the least privileged. Affirmative action, seen in this light, operates as a **non-ideal analogue** of the difference principle: it introduces inequalities (in the form of preferential treatment) that aim to benefit those historically excluded from fair competition.

While Rawls's ideal theory seeks to establish fair institutions from a neutral starting point, Thomson reminds us that **no real society begins in such a position**. The starting point itself is morally skewed by history. In such conditions, justice requires not impartiality in the abstract, but **partiality toward the wronged**—a deliberate rebalancing of the moral order.

In this sense, affirmative action embodies what Rawls might have called a **transitional principle of justice**: a set of non-ideal measures designed to bring society closer to the conditions that his ideal theory describes. It is a form of justice that acknowledges its own imperfection—an act of fairness performed within the limits of an unfair world.

Chapter 5

THE CAPABILITY APPROACH: SEN, NUSSBAUM, AND THE REFRAMING OF JUSTICE

5.1 Introduction

Amartya Sen's Capability Approach represents one of the most influential philosophical and practical re–orientations of justice in the last half–century. Sen — who combines development economics, welfare analysis and social choice theory (for which he received the Nobel Prize in Economics) — proposes a radically different way of thinking about human advantage and disadvantage. Instead of evaluating well-being by proxy indicators such as utility, income, or primary goods, Sen asks a more fundamental question: *what is a person actually able to do and to be?* Justice, on this account, is not a matter of maximizing utility nor merely equalizing resources, but of expanding the *real freedoms* of persons — their effective opportunities to live the lives they have reason to value.

The Capability Approach is genealogically rooted in multiple intellectual traditions. Sen often cites Aristotle's conception of *eudaimonia* — the idea that human flourishing is an activity, not a psychological state — as an early precursor. There are also strong connections to Adam Smith's theory of moral sentiments and his reflections on how social norms determine which goods count as "necessities," and to Marx's hope that just institutions should enable human beings to overcome the tyranny of circumstances. These influences converge in Sen's central thesis: individuals differ profoundly in their capacity to convert the same resources into functioning lives. "Equality of resources" is therefore too crude a metric. Real equality requires attention to human diversity.

Three conceptual commitments are essential:

- **Capabilities** — what a person can actually *do* with a given bundle of means. Two people with the same income may have radically different freedoms if one has a disability or lives under oppressive gender norms.
- **Adaptive preferences** — people internalize deprivation. The fox who cannot reach the grapes eventually says they are sour. Therefore, subjective preference data (and utilitarian "happiness" metrics) systematically underestimate injustice.
- **Freedom** — justice must measure both achieved functionings and the range of *genuinely valuable options* a person has, whether or not she exercises them.

In this respect Sen critiques both utilitarianism (too internal, too psychologistic, too insensitive to structure) and Rawls's "primary goods" metric (too resource-centered, insufficiently attentive to conversion factors). Our chapter will develop both lines of critique: how Sen seeks to re-think Rawlsian justice by replacing a measure of means with a measure of freedoms, and how this shift changes the evaluation of social arrangements — from domestic welfare analysis all the way into the global justice debate.

After presenting Sen's general framework and his distinctive conception of development as freedom, we will then turn to Martha Nussbaum, who develops a more normative, Aristotelian variant of the Capability Approach — one that proposes a list of central human capabilities as the political minimum required for a life worthy of human dignity. Whereas Sen resists such a canonical list (he insists capability evaluation should remain democratic, context-sensitive, and open-ended), Nussbaum argues that without a substantive threshold the approach cannot ground constitutional guarantees or human rights.

5.2 Sen's critique to utilitarianism

Sen's critique of utilitarianism reaches both the level of intuition *and* the level of formal axiomatic analysis. In his classic paper **"Utilitarianism and Welfarism" (Journal of Philosophy, 1979)** he demonstrated that utilitarianism is not simply a loose moral flavor or a general mood of "making people happy", but a *highly specific evaluative structure* built on two characteristic commitments:

1. **Welfarism** — all interpersonal comparisons must ultimately be reduced to information about individual utilities (understood as happiness, desire-satisfaction, or pleasure).
2. **Sum-ranking** — social evaluation is performed by maximizing the *aggregate* (or sometimes average) of these utilities.

Sen's argument is that the *combination* of welfarism and sum-ranking is normatively untenable. It forbids us, for example, from taking account of information about rights-violations, exploitation, vulnerability, unfreedom, or inequality *except* insofar as these things show up in utilities. And — this is crucial — they often do *not*. The deeply poor, the colonized, the subordinated, the socially excluded frequently adapt their self-perception and expectations to reduced circumstances. Their reported utility may look "fine". Their preferences may look "revealed". But their lack of capabilities is catastrophic.

This is why Sen says that utilitarianism fails as a public philosophy even when it "works" as a psychological description. His critique is not a marginal adjustment to classical welfare economics — it is a structural rejection of the evaluative foundations of utilitarian reasoning. Sen challenges each of the three pillars of utilitarianism: (1) the focus on mental states (pleasure, happiness, utility), (2) the sufficiency of revealed preferences, and (3) the reduction of human behavior to rational self-interest. For Sen, the problem is that utilitarianism measures “well-being” by what people *feel*, *desire*, or *prefer*, not by what they are *able* to do or to be.

Happiness vs. Pleasure.
Utilitarianism defines human well-being in terms of pleasure, happiness, or the satisfaction of desire. Sen argues that well-being depends not on feelings alone, but on *functionings* — the substantive things people manage to do or be: being healthy, being educated, participating in society, forming relationships, exercising practical reason. This Aristotelian insight — that eudaimonia is about worthwhile activity, not passive sensations — sits at the core of the Capability Approach. Pleasure is not irrelevant, but it cannot be the measure of human flourishing.

Democracy.
Utilitarianism treats political rights as instruments — useful only insofar as they increase net social utility. Sen, by contrast, insists that democratic rights have *intrinsic* importance. Democracy is not merely an aggregation mechanism or a voting technology; it is a practice of public reasoning and collective self-scrutiny. Famously, Sen showed that no functioning democracy with a free press has ever experienced famine — not because democracy maximizes utility, but because it institutionalizes accountability, information-flows, and open criticism.

Adaptive preferences.
Revealed preferences — the standard metric used by economists — cannot be taken as a reliable guide to well-being. People adapt to deprivation. They learn to “want less” when they are denied options. The fox that cannot reach the grapes declares them sour. A person long subjected to oppression may report being satisfied simply to avoid cognitive dissonance. Utilitarianism therefore risks *rewarding deprivation* — because the deprived often display low expectations and low reported dissatisfaction.

An objective theory of well-being.
Sen rejects the idea that subjective preferences, whatever they are, should determine what counts as well-being. “Sadistic” preferences or “expensive tastes” show how preference satisfaction can generate absurd or unjust results. Concepts such as exploitation, deprivation, urgency, vulnerability — all of which have deep roots in Marxian analysis — cannot be captured by subjective utility alone. They require objective evaluation of lives and opportunities.

Commitment.
Finally, Sen argues that human beings do not act solely out of self-interest. They act from commitment — moral identification with others, public causes, justice claims. Political action (protest, voting, solidarity) cannot be explained as utility maximization. The standard model of homo economicus is too thin to capture the kaleidoscopic nature of human motivation.

In short: Sen offers not merely a critique of Benthamite hedonism, but a re-founding of normative evaluation on a different basis — capabilities and substantive freedoms —

grounded in a richer anthropology of the human person, closer to Aristotle, Adam Smith, and (in some respects) Marx than to the standard economic calculus.

5.3 Sen's Critique of Rawls: Rethinking Rawlsian Justice in Capability Terms

Having shown why utilitarianism cannot serve as an adequate moral foundation — either empirically or normatively — Sen then turns his attention to Rawls. Rawls was, for Sen, the most important philosopher of justice of the twentieth century. But precisely because Rawls marks such a major advancement beyond utilitarianism, Sen insists that Rawls's theory must be interrogated and expanded further.

Rawls shifted the evaluative currency of justice away from utilities toward **primary goods** — rights, liberties, opportunities, income, wealth, and the social bases of self-respect. This was a decisive conceptual step and, to Sen's credit, a major breakthrough. Yet Sen argues that even primary goods remain an *indirect* metric. They are resources, not actual freedoms. The fact that someone possesses a given amount of goods "on paper" says nothing yet about what they can actually accomplish with those goods in their concrete circumstances.

This is where Sen's critique explicitly targets Rawls's informational basis. Sen criticizes Rawls's idea that primary goods are the appropriate measure for evaluating advantage. Think of a person in a wheelchair. Equalizing the amount of primary goods between an able-bodied and a handicapped person overlooks the fact that the latter would require more material and social resources to be mobile and to achieve other functioning, including overcoming social discrimination against disability. As Sen points out:

"The Difference Principle will give him neither more nor less on grounds of his being a cripple." In a later paper, he makes the more general point with even greater force: *"The primary goods approach seems to take little note of the diversity of human beings … If people were basically very similar, then an index of primary goods might be a good way of judging advantage. But, in fact, people seem to have very different needs varying with health, longevity, climatic conditions, location, work conditions, temperament, and even body size (affecting food and clothing requirements). So, what is involved is not merely ignoring a few hard cases [e.g. handicaps], but overlooking very widespread and real differences."* (Sen 1980)

This is the core problem. Individuals can convert resources into functioning with very different efficiency. A disabled person may require far more transportation resources to achieve basic mobility. A woman in a patriarchal context may be formally entitled to rights but socially unable to exercise them. A stigmatized minority may have legal equality but no real ability to transform this into participation or dignity. Equality of resources is therefore not equality of freedom.

Sen's conclusion is not that Rawls is wrong — but that Rawls's informational basis of justice is *too narrow*. Rawls remains one step inside the same paradigm of evaluation that Adam Smith, Marx, and Aristotle had already pointed beyond: that moral concern should attach not to goods but to the *substantive lives people are able to lead*. The

Capability Approach therefore presents itself not as an external adversary to Rawls — but as Rawls *in motion*. Sen universalizes Rawls's commitment to freedom by shifting the space of evaluation from **primary goods** to **capabilities**— from resources to *real opportunities*.

This has two systematic consequences.

(i) Capabilities are plural, not monistic.
Rawls still retains a single index of primary goods; Sen insists that human flourishing is multi-dimensional and cannot be reduced into one cardinal metric.

(ii) Capabilities are comparative across contexts.
Because conversion rates differ with social norms, technology, gender roles, disability, and cultural constraints, capability evaluation is unavoidably contextual and empirical — not merely theoretical.

In this sense, Sen's project returns political philosophy to the Aristotelian question: *what kinds of human lives are worth living?* And at the same time, his approach incorporates Marx's insistence on removing structural domination and Smith's insight that social respect requires material preconditions. Sen thereby synthesizes three traditions Rawls keeps mostly separate.

Ultimately, the Capability Approach is a generalization of Rawlsian justice into global space. Where Rawls's *Theory of Justice* was a domestic theory for a liberal nation-state, the Capability Approach immediately extends itself across borders — because capabilities are impaired by global institutions, and global justice must therefore be

5.4 Functionings and Capabilities

A central move in Sen's Capability Approach is the shift away from evaluating well-being in terms of resources, utility, or primary goods, toward evaluating what people are actually able to be and to do. This point may seem obvious, but it radically alters the informational space of justice. A bicycle, Sen notes, has the characteristic of being a means of transportation — but whether it actually provides transportation depends on the characteristics of the person who tries to use it. A bicycle means mobility for someone who is able-bodied, but for someone who cannot ride — either because of disability, or because the roads are unsafe for women to cycle — the physical possession of the bike does not translate into real mobility. What matters are not the goods as such, but the real opportunities they generate. This distinction allows Sen to introduce the core conceptual pair of the capability approach: functionings and capabilities.

Functionings refer to what a person actually manages to achieve — the states of being and doing that constitute a life. Being well nourished, having adequate shelter, being literate, participating in political processes, having stable self-respect — these are all examples of achieved states. Some functionings are elementary, like having enough food to survive. Some are more complex and socially mediated, like being able to appear with dignity in public, having a fulfilling job, or engaging in cultural or scientific activity.

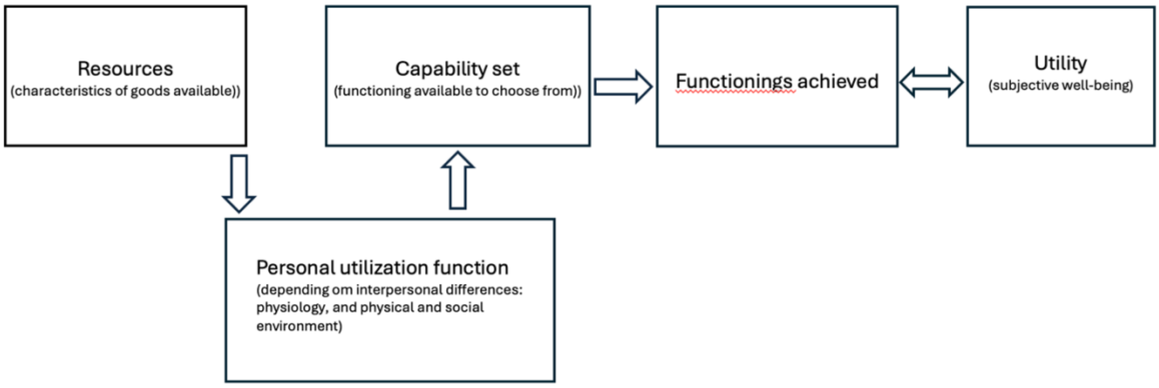


Fig. 5.1

Capabilities, in turn, refer to the real freedoms an individual has to achieve different functioning combinations — the set of valuable life paths a person could choose among. In other words, capabilities represent not what people in fact do, but what they are genuinely free to do. Two people can have the same functioning — for instance, both may be living on very little food — but their capabilities may differ dramatically. A person who is fasting out of religious conviction is choosing a functioning out of a wide capability set; the starving person in poverty has no such choice. The starving person manifests the same functioning state (not eating) but lacks the capability to eat. This distinction — between achieved states and genuine freedoms — is what makes the Capability Approach normatively richer than utilitarianism and also more discriminating than Rawls's focus on primary goods.

The capability perspective also illuminates a subtle but decisive point: what counts as necessary to live a dignified life is not fixed across time and space. It is socially and historically constituted. Adam Smith had already seen this in the eighteenth century when he observed that leather shoes and linen shirts had become "necessities" for social respect in England. Sen's capability approach generalizes this insight. In a rich, technologically advanced society, certain commodities — broadband access, transportation infrastructure, certain educational goods — are necessary to convert resources into functioning such as employability, civic participation, or cultural engagement. Poverty is relative in the realm of commodities but absolute in the realm of capabilities — because the capability space concerns what human beings can actually do and become, given the social context they inhabit.

Sen's distinction therefore reframes the question of justice: the object of moral concern should not be the mere possession of goods or the reporting of happiness, but the substantive freedoms to achieve lives people have reason to value. The capability approach is, in this precise sense, an Aristotelian, anti-welfarist, anti-resource-fetishist theory of justice. It measures advantage not by satisfaction, not by possession, but by the real, empirically traceable space of human possibility.

The capability approach therefore requires us to analyze not only what commodities people possess, but how those commodities translate into real functionings across different persons and different contexts. The conversion of commodities into capabilities is socially and biophysically mediated. First, individual physiology matters. Two persons with the same bundle of goods may not enjoy the same level of functioning, because the requirements for achieving a given functioning vary with bodily features. A person with a chronic illness may need expensive medication and more nutritional intake simply to achieve a basic level of health; a disabled person may require assistive technologies to obtain mobility. Equal resources do not imply equal capability because the cost of transforming commodities into functioning varies across bodies.

Second, local environmental diversities impose different conversion burdens. Climate affects heating and clothing needs; pollution affects respiratory health; epidemiological profiles alter the required expenditure on medical resources. A household in a malaria-endemic region may require mosquito-nets, anti-malarial treatments and public sanitation investments merely to maintain a minimal functioning threshold that a household in a temperate, low-disease environment achieves without effort. The same basket of goods generates different real opportunities depending on where one lives — a point utilitarianism and Rawls's primary goods metric leave analytically underdeveloped.

Third, there are variations in social conditions, especially with respect to the provision of public goods. Safe transport, schooling, roads, legal protections, policing, digital infrastructure — all alter the capability to use commodities effectively. A laptop is not a capability for online education where there is no broadband; the right to vote is not a functioning when intimidation or clientelism prevent one from exercising it. Community relationships also matter. Dense social networks can reduce vulnerability and expand effective choice sets, just as social isolation can shrink them.

Fourth, the capability approach takes seriously the role of relational norms and social expectations. What counts as a "necessary" commodity is partly determined by conventions of dignity. Smith's insight about linen shirts and leather shoes applies: social respectability has commodity preconditions. The inability to meet socially embedded standards of public appearance can render persons ashamed, humiliated, excluded — even when they are technically not starving. Capability disadvantage is therefore partly relational: it concerns one's standing among others and the demands of social practice.

Fifth, Sen emphasizes that capability evaluation must look inside the household. Resources allocated to a family are not automatically converted into capabilities for each member. Distributional rules often favor males over females, adults over children. This means that household income is not a reliable proxy for individual capability. A girl in a patriarchal household may have worse real opportunities than a boy in a household with lower average income. The capability approach therefore disrupts the assumption that the household is a unified decision-making unit. It insists that justice must be evaluated person-by-person, not household-by-household.

Taken together, these five conversion differences show why the move from resources to capabilities is not a philosophical luxury but an empirical necessity. The capability approach is grounded in the way the world actually works. It explains why two individuals with identical bundles of goods can lead radically unequal lives, and why justice must be assessed not by what people have, nor by what they feel, but by what they can in fact do and become.

5.5 Agency vs Well-Being — Beyond Utilitarian Beneficence

One of the most important conceptual shifts introduced by Sen is the sharp — but analytically fertile — distinction between *well-being* and *agency*. Classical utilitarianism, and even much contemporary humanitarian ethics (for example in Peter Singer's global beneficence approach), treat persons primarily as *patients* — as passive recipients of utility, welfare, pleasure, or benefit. In these frameworks, the central question is: *how can we maximize the well-being of individuals?* and the answer is conceived predominantly in terms of allocating resources or acting on behalf of those whose utility we want to raise.

Sen does not deny the moral importance of well-being. But he argues that to see human beings *primarily* as patients is an ontological and ethical reduction. Human beings are also — and fundamentally — *agents*: we pursue projects, set goals, intervene in the world, form commitments, take responsibility for what we do, and care about justice or reasons that are not confined to our own utility. Agency is thus the capacity to act on values, not simply to *receive* welfare.

And crucially, Sen stresses that enhancing well-being is not independent of agency. In most real social contexts — development, poverty policy, women's empowerment, education reform — the most effective way to improve people's well-being is not to merely "give them more utility", but to expand their *capacity to shape their own lives*. Agency — participation, voice, political empowerment, the ability to make claims and shape institutions — is a route *to* well-being. It is therefore not an optional add-on or a secondary moral category: it is constitutive of what it means to treat human beings as ends in themselves.

This is where Sen breaks decisively from utilitarianism — and also from Singer's model of benevolent benefactor morality. Singer's demandingness shows that utilitarianism can produce radical duties — but it still positions the poor or distant needy primarily as *objects* of aid: people whose well-being we should maximize. Sen believes this misses the point. The normative focus should not be on maximizing the welfare of passive recipients — but on enabling *capable agents* to actively determine their own futures. Agency is not simply instrumentally valuable for its contribution to utility; it is part of human flourishing itself.

In this sense, the capability approach is not only an evaluative theory — it is also an ethical anthropology. It builds a picture of the human person as a doer, a chooser, a participant, a co-author of their social world. And therefore justice, on this view, is not measured by how much welfare flows into people's lives — but by whether the social world equips people with the substantive freedoms to shape their lives *as agents*. This, for Sen, is the deep moral break with utilitarianism: human dignity lies not in how well we are treated, but in our capacity to take part in determining the shape of our own life.

5.6 Freedom and Development

For Sen, the central organizing thesis of his mature work — particularly in *Development as Freedom* (1999) — is that **development is a process of expanding human freedom**, not merely a process of increasing income, industrial output, or GDP. This is not a poetic metaphor: it is a radical reframing of what counts as progress. Development, Sen argues, cannot be evaluated by commodities, resources, or utility alone. Rather, a society is more developed *insofar as* its people have more substantive opportunities to live the kinds of lives they have reason to value.

This is why the capability approach is inseparable from a theory of freedom. Capabilities *are* freedoms — but not the purely formal liberties of classical liberalism (freedom of speech, freedom of contract) nor the purely "negative" freedoms of non-interference. They are **effective freedoms**: what a person is actually able to do and to be, given her social position, her health, the norms of her community, and the institutional environment. Two individuals may each have the formal legal right to run for office — but only one may have the education, the financial autonomy, and the social respect needed to realistically do so. Formally equal liberties, without support, do not yield real capability.

Thus Sen sees a continuity between individual agency and collective development. Freedom is at once the *means* and the *end* of development. When people can participate in political life, when women have property rights and literacy, when laborers are not dependent on local elites, when minorities have voice and legal protection — these "freedom expansions" do not merely increase utility; they *cause* material progress. Famines do not occur in functioning democracies, Sen famously argued, because political freedoms generate public scrutiny, accountability, information flows, and pressure for action. Freedom creates development — not only development enabling freedom.

This also implies that institutions must themselves be judged not by aggregate outcomes but by their **impact on expanding people's capabilities**. Markets, for instance, can be capability-enhancing — but only under conditions that protect people's basic capabilities, provide public goods (education, health, security), and counter structural forms of exclusion. A market that increases average income while deepening capability deprivation for women or low-castes or the disabled is not, on Sen's terms, "successful".

Freedom, therefore, is not just an ideal of personal autonomy: it is the **metric and mechanism of justice**. A society is not more developed when it has more wealth — it is more developed when it has less unfreedom. And this is the crucial normative difference between Sen's view and utilitarian or resource-based approaches. Development is not the maximization of welfare *from outside*, but the empowerment of people *from within*. It is not the benevolent redistribution of goods to patients, but the expansion of agency — enabling human beings to lead, in Aristotle's phrase, "a life fit for a human being".

The result is a theory of progress that is both empirically grounded and normatively ambitious. Sen replaces the question "how much income do people have?" with the more philosophical — but ultimately more realistic — question: **what can people actually do?**
From this vantage point, poverty, inequality and injustice are not merely deficits of resources or happiness — they are deficits of freedom. And development is the long, incomplete, political work of removing those unfreedoms, so that human agency has room to flourish.

Sen's focus on freedom also places him at a decisive crossroads in the contemporary debate on global justice. Rawls, in *The Law of Peoples*, maintains a state-centered architecture: justice applies primarily *within* well-ordered societies, and the global domain is regulated mainly by *duties of assistance* between "peoples", not by distributive egalitarian principles among *individuals*. In Rawls's view the primary task internationally is enabling "burdened societies" to become "well-ordered"; once this threshold is reached, inequality between societies need not trigger further duties of redistribution. Sen departs here. Because capabilities are *individual* freedoms — and because global institutions dramatically shape the space of individuals' opportunities — the global arena cannot be treated merely as a system of cooperation between states. In other words: capability deprivation is experienced by persons, not by "peoples".

This is precisely where Pogge and Beitz break more radically with Rawls, each in his own way. Pogge argues that the current global order itself actively *harms* the global poor (for example via resource-privilege and borrowing-privilege rules that empower oppressive domestic elites). Therefore, global justice is not primarily a charitable "duty of assistance", but a duty *not to impose institutional structures that foreseeably perpetuate severe capability deprivation*. Beitz, earlier still, argued that if distributive justice is appropriate within a state because economic interdependence and shared institutions create a system of cooperation, then — given extensive global interdependence — distributive principles must extend *beyond* borders as well.

Sen stands between Rawls and the cosmopolitans. He agrees with Rawls that freedom (not utility) is the proper space of moral concern, but rejects the Rawlsian restriction of primary moral equality to the domestic basic structure. Sen does not argue — like Pogge — that the global order is necessarily harmful in every dimension, nor — like early Beitz — that Rawlsian egalitarianism must simply be "scaled up" globally. Instead, Sen reframes the entire question: global justice is not anchored in a *unique* principle (equality, difference principle, utility) but in a comparative evaluation of capability deprivation wherever it occurs — domestically or internationally. There is no sharp boundary between domestic and global spheres when the relevant unit of concern is *the human being as agent*.

Thus Sen subtly recasts the Rawls–Pogge–Beitz debate: he is not primarily extending, or negating, Rawls — he is shifting the normative vocabulary from *resources* to *freedoms*, from "duties of peoples" to the lived capabilities of persons. This move is precisely what allows the capability framework to become, in the next step (Nussbaum, global capability lists), a philosophical architecture suitable to a **Philosophy for Global Justice** — where justice is evaluated in the widest possible space: *the real lives that human beings are actually able to live, across borders*.

5.7 Nussbaum: From Cosmopolitanism to the Capability Approach

Martha Nussbaum enters the global justice debate originally through the route of cosmopolitanism. From her earliest writings — for example her commentary on the Stoics and Cicero — Nussbaum argues that political morality must take the individual human being, not the nation or the people, as the primary unit of concern. This is the classical cosmopolitan thesis: *I am a citizen of the world first,* and only derivatively of Athens, Rome, India, or the United States. Borders do not morally constrain the scope of our duties.

This foundational intuition places her immediately in tension with Rawls's *Law of Peoples*, which treats organized "peoples" as the entities of international justice. Nussbaum argues instead that it is the *person* — not the *people* — that is owed justice. This shift in the basic subject of justice is the bridge that leads her from cosmopolitanism to the Capability Approach.

Nussbaum adopts Sen's core insight — that what matters morally is not the resources a person has but what they are *actually able to do and to be*. But where Sen leaves the capability space formally open, Nussbaum believes that a serious theory of justice must specify a determinate list of capabilities that are necessary conditions of a fully dignified human life. For her, human dignity is not merely a procedural or comparative standard — it is a substantive Aristotelian claim about what flourishing consists in. Dignity therefore requires thresholds. If a political order leaves persons below the threshold in bodily security, bodily health, practical reason, political participation, etc., it is not merely sub-optimal — it is unjust.

Thus: Nussbaum *cosmopolitanizes* the capability approach. She proposes that each person on earth has a claim of justice to a set of core capabilities, and that these claims are *not* conditional on membership in a political community, nor on reciprocal cooperation, nor on contributing to a "people." In contrast to Rawls, Nussbaum does not restrict global justice to a duty of assistance; she treats the global order itself as subject to the same moral requirements as the domestic constitution. States, international organizations, and global regimes are ethically legitimate only if they protect and secure the central capabilities for *every* person.

In short:

- Cosmopolitanism gives Nussbaum her *unit*: the individual person.
- Aristotelian dignity gives her the *normative foundation*: the flourishing life.
- The Capability Approach gives her the *metric*: the threshold of substantive opportunities.

The result is the first major fully cosmopolitan theory of the capabilities tradition: a framework which insists that justice is owed across borders, not only within them — and that the measure of global justice is the extent to which every human being, anywhere on earth, can genuinely develop and exercise the capabilities necessary for a life worthy of human dignity.

Nussbaum's "Ten Capabilities"

Martha Nussbaum's contribution to the capability approach represents a decisive shift from methodological openness to explicit normative prescription. While she fully shares Amartya Sen's rejection of utility-based theories of well-being, and his insistence that justice must evaluate what people are actually able to *be* and *do*, she argues that Sen ultimately does not go far enough. Sen wants freedom to be understood as an open-textured space of possibilities. Nussbaum insists that this space must be anchored by substantive, universal, cross-culturally defensible claims about human dignity. Her work thus takes the capability approach from its original role as an evaluative space — a way of comparing lives — to a fully articulated theory of justice grounded in moral philosophy, human rights, and cosmopolitanism.

Nussbaum draws directly on Stoicism, Enlightenment universalism, and contemporary feminist theory to argue that human dignity is not a local cultural construct, but a moral fact about the kind of beings we are: we are vulnerable, embodied, needy creatures who reason, who seek affiliation, who depend on care and who have a claim to lives worthy of that human status. For this reason, the capability approach must specify certain entitlements as non-negotiable conditions of justice. Unlike Sen, Nussbaum therefore presents a list — not as a sociological description of what nations tend to value, but as a normative threshold that every person everywhere has the right to reach. These are the "central capabilities," and they form the core of her version of global justice.

1. **Life** – Being able to live to the end of a normal human lifespan; not dying prematurely.
2. **Bodily Health** – Being able to have good health, adequate nutrition, shelter, reproductive health.
3. **Bodily Integrity** – Being able to move freely, to be secure against violence (domestic and public), and have sexual autonomy.
4. **Senses, Imagination, and Thought** – Being able to use one's mind and senses, informed by adequate education; freedom of expression; enjoyment of the arts.
5. **Emotions** – Being able to love, form attachments, and not have one's emotional development damaged by fear, trauma, humiliation.

6. **Practical Reason** – Being able to form a conception of the good and engage in critical reflection; freedom of conscience.
7. **Affiliation**
   – A: Being able to live with and toward others with respect and non-humiliation
   – B: Having the social bases of self-respect; protection against discrimination.
8. **Other Species** – Being able to live with concern for animals, plants, nature.
9. **Play** – Being able to laugh, play, enjoy recreation.
10. **Control over One's Environment**
   – A: *Political* – Being able to participate effectively in political choices.
   – B: *Material* – Being able to hold property and work as a participant in the labor market on equal terms.

Each capability on the list indicates not just a functioning (for example, actually participating politically), but a *real freedom* to choose the functioning (for example, the genuine ability to speak, organize, and vote). In this way the list rejects both paternalism (because people may legitimately choose not to exercise a capability) and relativism (because the capability must still be *there* as a real option).

In the book *Creating Capabilities* Nussbaum develops two central ideas around this list.

First, justice must be *person-by-person*. Aggregate indicators — GNP, average utility, average literacy — mask relational inequality and the silencing of vulnerable groups. Distributive justice must see each person as an end, not as a statistical unit.

Second, the capability threshold is a matter of *urgent* moral priority. Nussbaum's approach is not about maximizing total capability in the world but securing a floor: a moral minimum required for a life worthy of human dignity. Only after this floor is secured can legitimate differences in culture, lifestyle, or market distribution unfold.

For this reason, Nussbaum's theory is explicitly cosmopolitan: each person is a holder of a claim that extends beyond nation-state borders. The unit of concern is not "my compatriot" but the human person. Where Sen had opened a conceptual space for thinking about global justice, Nussbaum turns that space into a normative architecture — a set of rights-like entitlements that every political system ought to secure.

Thus, Nussbaum's capability approach is a natural successor to — and an expansion of — Sen's. Where Sen pluralizes the informational space of justice, Nussbaum anchors that plurality in a universal moral core. In doing so, she offers one of the most influential philosophical proposals for a 21st-century cosmopolitan theory of global justice.

***Why this list matters in* Creating Capabilities**

Nussbaum's claim is that these ten capabilities are not simply a philosophical exercise, nor a rough guide for policy analysts — they represent the *minimum* conditions of social justice. A society that fails to secure these central capabilities for all its members is, on Nussbaum's account, not merely sub-optimal: it is *unjust*. The list therefore has a status comparable to a constitutional bill of rights — but one grounded not in tradition or national history, but in an account of human dignity. Each capability identifies a distinct dimension of what it means to live a life worthy of human beings, and each is non-fungible. We cannot compensate for the violation of one capability (e.g. bodily integrity) by providing excess in another (e.g. access to education). This *plurality and incommensurability* are a central departure from utilitarian and welfarist approaches.

Nussbaum characterizes the list in three key ways:

universal but not culturally imperialist — because they set thresholds, not a complete way of life
The list does *not* prescribe a single conception of the good life (e.g. Western secular individualism). Instead, it establishes a minimum floor — a threshold — of substantive freedom beneath which no person should fall, regardless of culture. Societies remain free to construct and interpret flourishing beyond that floor. Capabilities require that one *can* choose certain functionings, not that one *must*. This is what makes the list universal without imposing a homogenized worldview. Muslim, Buddhist, Indigenous, Christian, secular — each can interpret, within their cultural horizon, what a flourishing life looks like *beyond* the threshold.

grounded in dignity rather than utility or resources
The list reflects the conviction that human life has a worth that cannot be reduced to pleasure (utilitarianism), or to bundles of commodities (resourcism), or to Rawlsian primary goods. What matters is what a person can *be and do*. Dignity attaches to persons as ends, not as containers of satisfaction or as vectors of purchasing power.

suitable for constitutional protection and policy evaluation
Because the list does not merely measure outcomes — but identifies the normative baseline that defines justice — it can serve as a basis for real political institutions. Nussbaum explicitly proposes that capabilities should be written into constitutions as positive entitlements. The list is therefore not only philosophical: it is a *criterion for judging whether governments are doing their job*.

Nussbaum's book defends four core theses, and she defends them not only philosophically, but with public-reason arguments that can travel across traditions:

1. **Development is freedom and dignity — not GDP.**
   Economic growth is instrumentally useful — but irrelevant in itself. What matters is whether growth expands people's substantive opportunities. A country may become richer while restricting women's mobility, excluding disabled citizens, or destroying ecosystems. GDP cannot tell us whether people can live lives they have reason to value — capabilities can.
2. **Justice requires guaranteed substantive opportunities (not merely formal rights).**
   Formal rights without actual capability are often fictitious. A woman who "has the right to vote" but cannot leave the house without male permission does not possess political capability. A disabled person who "has the right to work" but

lacks ramps or assistive devices does not have employment capability. Justice is about *real freedom*, not merely paper-freedom.

3. **Capabilities — not resources — should be the metric of social evaluation.** Resources can be distributed equally while freedoms remain radically unequal. The same bicycle gives mobility to one person and zero mobility to another (if she has a spinal cord injury). The relevant moral question is not "what do you have?" but "what can you *do* with what you have?" Capabilities register this conversion space — the zone between commodities and real opportunity.
4. **These claims hold globally, not just within national borders.** Justice cannot be contained within the nation-state. The very same capability thresholds apply to a woman in Lagos, a farmer in Bihar, or a pensioner in Minnesota. If dignity is universal, then political obligation cannot stop at the border. *Creating Capabilities* therefore pushes the capability approach into cosmopolitan territory — a global architecture of responsibility in which every person, not every "citizen," is the moral unit of analysis.

This positioning of Nussbaum's list completes the conceptual shift from Sen's evaluative space to a full-fledged normative theory of global justice — one in which human dignity is the ultimate currency, capabilities are the metric, and the world — not the nation — is the scope.

Nussbaum's claim is that these ten capabilities are not simply a philosophical exercise, nor a rough guide for policy analysts — they represent the *minimum* conditions of social justice. A society that fails to secure these central capabilities for all its members is, on Nussbaum's account, not merely sub-optimal: it is *unjust*. The list therefore has a status comparable to a constitutional bill of rights — but one grounded not in tradition or national history, but in an account of human dignity. Each capability identifies a distinct dimension of what it means to live a life worthy of human beings, and each is non-fungible. We cannot compensate for the violation of one capability (e.g. bodily integrity) by providing excess in another (e.g. access to education). This *plurality and incommensurability* are a central departure from utilitarian and welfarist approaches.

Nussbaum characterizes the list in three key ways:

universal but not culturally imperialist — because they set thresholds, not a complete way of life
The list does *not* prescribe a single conception of the good life (e.g. Western secular individualism). Instead, it establishes a minimum floor — a threshold — of substantive freedom beneath which no person should fall, regardless of culture. Societies remain free to construct and interpret flourishing beyond that floor. Capabilities require that one *can* choose certain functionings, not that one *must*. This is what makes the list universal without imposing a homogenised worldview. Muslim, Buddhist, Indigenous, Christian, secular — each can interpret, within their cultural horizon, what a flourishing life looks like *beyond* the threshold.

grounded in dignity rather than utility or resources
The list reflects the conviction that human life has a worth that cannot be reduced to

pleasure (utilitarianism), or to bundles of commodities (resourcism), or to Rawlsian primary goods. What matters is what a person can *be and do*. Dignity attaches to persons as ends, not as containers of satisfaction or as vectors of purchasing power.

suitable for constitutional protection and policy evaluation
Because the list does not merely measure outcomes — but identifies the normative baseline that defines justice — it can serve as a basis for real political institutions. Nussbaum explicitly proposes that capabilities should be written into constitutions as positive entitlements. The list is therefore not only philosophical: it is a *criterion for judging whether governments are doing their job*.

Nussbaum's book defends four core theses, and she defends them not only philosophically, but with public-reason arguments that can travel across traditions:

1. **Development is freedom and dignity — not GDP.**
   Economic growth is instrumentally useful — but irrelevant in itself. What matters is whether growth expands people's substantive opportunities. A country may become richer while restricting women's mobility, excluding disabled citizens, or destroying ecosystems. GDP cannot tell us whether people can live lives they have reason to value — capabilities can.
2. **Justice requires guaranteed substantive opportunities (not merely formal rights).**
   Formal rights without actual capability are often fictitious. A woman who "has the right to vote" but cannot leave the house without male permission does not possess political capability. A disabled person who "has the right to work" but lacks ramps or assistive devices does not have employment capability. Justice is about *real freedom*, not merely paper-freedom.
3. **Capabilities — not resources — should be the metric of social evaluation.**
   Resources can be distributed equally while freedoms remain radically unequal. The same bicycle gives mobility to one person and zero mobility to another (if she has a spinal cord injury). The relevant moral question is not "what do you have?" but "what can you *do* with what you have?" Capabilities register this conversion space — the zone between commodities and real opportunity.
4. **These claims hold globally, not just within national borders.**
   Justice cannot be contained within the nation-state. The very same capability thresholds apply to a woman in Lagos, a farmer in Bihar, or a pensioner in Minnesota. If dignity is universal, then political obligation cannot stop at the border. *Creating Capabilities* therefore pushes the capability approach into cosmopolitan territory — a global architecture of responsibility in which every person, not every "citizen," is the moral unit of analysis.

-

This positioning of Nussbaum's list completes the conceptual shift from Sen's evaluative space to a full-fledged normative theory of global justice — one in which human dignity is the ultimate currency, capabilities are the metric, and the world — not the nation — is the scope.

5.8 Policy implications: from theory to human development practice

Sen's Capability Approach has not remained a purely philosophical construct; it has exercised direct influence on the priorities, metrics and discourse of global development policy. Its most visible institutional translation is the Human Development paradigm inaugurated by the United Nations Development Programme in 1990. Motivated explicitly by Sen's claim that development is "expansion of capabilities" rather than mere growth, the Human Development Reports (HDR) redefined what it means for a society to advance: not "how much output", but "how much real opportunity to live worthwhile lives". In operational terms, the HDR introduced the Human Development Index (HDI), which approximates three core capability domains: the capability to live a long life (life expectancy), the capability to become literate and communicate (education), and command over resources to pursue one's ends (via GDP per capita). None of these is a direct measure of pleasure or utility; they are functioning-based proxies for freedom to achieve. The HDI therefore illustrates — in statistical form — the central theoretical shift Sen had argued for: that capability-based evaluation yields a radically different picture of cross-national justice than resource-based or income-based metrics. Countries with high GDP but low literacy or low life expectancy fall dramatically in rank; some with modest incomes but robust education and public health climb sharply. This is precisely Sen's normative point: well-being cannot be read through commodities alone.

Here Nussbaum's contribution gradually enters the policy space as well. While Sen refuses to specify a canonical list of central capabilities, Nussbaum offers a philosophically grounded, explicitly normative list of ten "endowments of dignity" that she believes states have a duty to secure at least to a threshold level. Her work in *Creating Capabilities* has influenced gender policy frameworks, constitutional design debates in India and South Africa, and the evaluative architecture of disability-policy in the OECD. Whereas Sen's framework provides the methodological pluralism and evaluative caution the HDR uses, Nussbaum's list provides the moral minimalism for what counts as a life worthy of human dignity. In practice, one can say: the HDI is "Sen-inspired measurement", while threshold-capability constitutionalism is "Nussbaum-inspired institutionalisation".

Taken together, Sen and Nussbaum thus mark a distinctive normative break with both utilitarian welfareism and Rawlsian resource-based egalitarianism within public policy. Instead of asking *how much income* a person has — or *how satisfied* she reports being — capability-based policy asks: *what can she actually be and do?* The implication is structural and radical: the state is responsible not merely for formal rights, but for the enabling conditions — education, healthcare, gender equality, bodily integrity, democratic voice — that allow those rights to become real. This is where their work directly engages global justice: if capabilities are the proper space of moral evaluation, then capability deprivations that are the result of global institutional arrangements (trade rules, IP regimes, global finance rules, climate asymmetries) are not merely unfortunate — they are injustices. Policy, on this view, is not redistributive charity; it is structural repair.

Chapter 6

COMMUNITARIANISM

6.1 Introduction

Communitarianism is not a homogeneous cultural or philosophical movement. Thinkers often grouped under this label — Charles Taylor, Michael Sandel, Michael Walzer, among others — are united less by a common positive doctrine than by a shared *critique* of liberalism. What makes them "communitarian" is not a unified systematic position, but a family resemblance: they contest the liberal assumption that the self is, in essence, a free and unencumbered chooser whose primary relation to society is instrumental.

Modern communitarianism, in its contemporary form, begins as a philosophical reaction to John Rawls's *A Theory of Justice*. Drawing in different ways on Aristotelian resources, communitarian thinkers challenge the liberal idea that the main task of government is primarily to secure and distribute fairly a set of rights and resources so that individuals may pursue freely chosen "life projects". Against this view, communitarians insist that our ethical and political agency is not prior to social relations, but formed by them.

Their critique of liberalism has been articulated along three main lines:

1. **The moral importance of traditions and social contexts** is underestimated by liberal individualism.
2. **The self is misconceived** when imagined as an abstract chooser, independent of historical and communal attachments.
3. **Communal membership and shared practices** are not incidental external relations — they are constitutive of agency and value.

From a communitarian perspective, principles of justice cannot be derived from abstract hypothetical devices such as Rawls's "original position". Political philosophy must instead begin from the habits, histories, practices and interpretive traditions of actual people living in specific places and times. Liberalism cannot legitimately claim universal moral jurisdiction.

The political optimism of the immediate post–Cold War moment — when many believed (following Fukuyama's "end of history" thesis) that liberal democracy would become globally hegemonic — has faded. Ethnic conflict, persistent poverty, entrenched corruption, and institutional fragility have shown that the transplantation of liberal institutions cannot be taken for granted.

Yet the deepest theoretical challenge to Western liberal democracy has not only come from post-Soviet disillusionment; it has also emerged from East Asia. Asian political cultures often place special emphasis on family, social harmony, and the precedence of collective interests over individual choice. China's rise — economically, culturally, and politically — has made this challenge geopolitically tangible.

Here, communitarian arguments take a distinctly *cultural* form — a form of cultural particularism. Three central claims emerge:

1. **Cultural factors affect the prioritization of rights.**
   American citizens may be more willing to sacrifice social rights in order to secure political liberties; Chinese citizens may priorities social and economic welfare even at the cost of individual freedoms. Confucian cultures allocate resources to education because of inherited ethical valuations, not only because of economic calculus.
2. **Cultural factors affect the justification of rights.**
   In Singapore, for instance, democratic rights may be defended not by appealing to individual freedom, but because they strengthen ties to communities such as the family or the nation.
3. **Cultural factors provide moral foundations for distinctive political institutions.**
   Confucian ethics emphasize filial piety. This raises contemporary political questions: should the duty to care for elderly parents be enforced by legal mandate, or realized through tax incentives? Some Confucian theorists even propose an additional chamber of government — a "House of Exemplary Persons".

At the heart of all of this is the communitarian claim that Rawlsian liberalism rests on an overly individualistic conception of the self. When Rawls assumes that our highest interest is to pursue our chosen life plan, he overlooks the way in which our identities are constituted by familial, cultural, and historical attachments. Politics, communitarians argue, should not only secure conditions for individual choice; it should also cultivate and sustain the forms of social belonging that make meaningful agency possible.

The goods to which we are committed are not invented ex nihilo by self-creating individuals. They arise from communities, practices, languages, and traditions. Choice, in itself, is not intrinsically valuable: we do not choose our family, nation, or history — and we do not relate to them *as if* they were mere objects of option. They are the background through which *value*, *identity*, and *moral understanding* become intelligible at all.

This is the starting point of the contemporary communitarian critique.

Communitarianism and Global Justice

The communitarian critique has important implications for debates on global justice. If moral reasoning is embedded in particular histories and social practices, then the idea of a single universal set of global principles becomes problematic. Cosmopolitan theories of global justice — whether derived from Rawls, or developed in contemporary form by thinkers like Beitz and Pogge — assume that there exists, in principle, a standpoint from which justice for all humanity can be defined independently of any specific culture. Communitarianism denies precisely this assumption. Justice, communitarians argue, is not a free-floating rational construction. It is generated within communities, languages, and traditions. Without a shared cultural horizon, the very meaning of fairness and of the good life becomes indeterminate.

This is why, for example, Michael Walzer rejects the idea of a thick, substantive theory of global distributive justice. For Walzer, we can only have very "thin" moral agreements at the international level — basic norms prohibiting genocide, slavery, or mass starvation — because genuine "thick" justice presupposes shared meanings that only exist inside historically constituted political communities. In this sense, communitarianism functions as a constraint upon universalistic projects. It does not say that global justice is impossible, but that it cannot be defined through abstract principles alone. It must grow from concrete historical forms of life.

Furthermore, cultural variants of communitarianism — as developed in contemporary Confucian political thought — reinforce this conclusion from a different angle. East Asian societies with communitarian values do not always give priority to individual rights in the same lexicographic order as Western liberal democracies. Social harmony, filial piety, the prosperity of the family, or the stability of the community may be regarded as ethically prior to individual autonomy. This suggests that there may not be one single model of just social order for all societies, but several legitimate pathways. Communitarianism therefore does not only challenge liberal accounts of the self: it also challenges the aspiration to a single, universalistic conception of global justice.

In short, the communitarian position shifts the debate from universalism to pluralism. It invites us to see global political philosophy not as the search for one set of principles valid everywhere, but as a negotiation between culturally situated moral worlds.

6.2 Charles Taylor: The Critique of Atomism and the Principle of Belonging

Charles Taylor is a Canadian philosopher whose work has had a profound influence on contemporary debates about the modern self. In *Sources of the Self: The Making of the Modern Identity*, he reconstructs the genealogy of the Western self, showing that it has been shaped by multiple historical strands: the affirmation of individual freedom; the ideal of personal exploration; the sense of nature as a source of moral meaning; the valuation of authenticity; the ethical dignity of ordinary life; and a moral pull toward benevolence and solidarity with the suffering of others. Taylor's broader project is hermeneutical: to understand social life, one must understand the self-interpretations of the agents who take part in it. Social science cannot be reduced to causal modelling of interest-maximising behaviour. Voting, for example, is not only strategic choice — it is an expression of democratic identity, of civic belonging, of shared concerns and shared meanings.

Taylor is commonly labelled a "communitarian" because he stresses the social nature of the self and the obligations we owe to the communities in which we are embedded. His central target is what he calls **atomism**. Atomism refers to doctrines associated with the social contract tradition — Hobbes, Locke, and more recently Nozick and Rothbard — doctrines which assume that individuals and their rights are prior to society, and that social order is merely instrumental. Neo-liberalism is the contemporary political expression of this anthropological thesis. Atomists defend a primacy of rights while denying an equally fundamental principle of belonging.

Taylor contrasts this with the Aristotelian thesis that human beings are "political animals". On this classical view, individuals do not develop their capacities *outside* society; they become fully rational, responsible, moral agents *only as members* of a polis. Atomism reverses this relation: it asserts that the individual is self-sufficient and only contingently social.

A liberal might reply that this is irrelevant. We grant rights even to those who currently cannot exercise rational capacities — infants, the severely disabled, people in coma — therefore rights cannot depend on actualised capacities. But this objection misses the deeper point. When we ascribe rights to persons we do not simply prohibit certain harms; we implicitly affirm that these beings *are the kinds of beings* whose flourishing matters — because they possess the potential for certain characteristically human capacities. Even animal rights debates show this structure: to insist that sentient animals deserve moral consideration is to say that their capacity to suffer, feel pleasure, and pursue a good of their own gives rise to normative claims.

Thus: rights presuppose a picture of what is worthy of respect — a picture of human capacities.

Once this is recognised, Taylor argues, the liberal doctrine of the primacy of rights becomes incoherent if it claims to be anthropologically neutral. If rights are important because the distinctively human capacities are important, then we are committed not only to recognising rights but also to **promoting the development of those capacities**. And those capacities do not develop in isolation. They develop only in certain social contexts — contexts that sustain language, recognition, education, practices of discourse, and shared moral horizons.

For this reason, Taylor concludes that the assertion of rights necessarily implies an **obligation to belong**. The liberal self could not exist without the social conditions that form it. Individual rights matter to us because the characteristically human capacities matter to us; if these capacities matter, then the social forms that sustain and nurture them must matter as well.

Belonging, therefore, is not an optional addition to liberal rights.
It is conceptually inseparable from them.

Taylor's argument against atomism opens the space for a broader critique of liberalism understood as a doctrine of the unencumbered self. But Taylor is not alone in this line of thought. In the American context, Michael Sandel develops a parallel challenge to Rawls from within the analytic tradition. Whereas Taylor proceeds genealogically and hermeneutically, Sandel engages Rawls more directly on conceptual grounds. Yet both converge on the same core point: if the self is constituted by prior attachments, then the Rawlsian image of an agent who can step outside all contingencies and choose principles of justice from a neutral standpoint is incoherent. Sandel radicalises this insight by arguing that the very structure of Rawls's "original position" presupposes what it seeks to deny — a pre-social, autonomous chooser. It is to this critique that we now turn.

6.3 Michael Sandel: Liberalism, Moral Neutrality, and the Limits of the Unencumbered Self

Michael Sandel endorses a form of civic republicanism or communitarianism — although he is not fond of such labels. His first major book, *Liberalism and the Limits of Justice* (1982), is a direct challenge to Rawls's *A Theory of Justice*. Rawls attempts to provide a universal, value-neutral justification for liberalism based on the priority of the right over the good: the principles of justice, according to this approach, can be defined independently of any particular moral, cultural, or religious conception of the good life. Sandel argues that this supposed moral neutrality is not only unachievable in practice — it is conceptually incoherent.

For Sandel, we cannot resolve political controversies about abortion, same-sex marriage, free speech, slavery, or human dignity solely by appealing to neutral rights-language. These disputes inevitably involve substantive moral judgments. Therefore, meaningful democratic deliberation cannot avoid moral and religious argument — because what is at stake is precisely the question of which goods deserve respect, protection, or institutional support.

Sandel's deeper objection is that Rawls presupposes a certain *theory of the self*. Rawls must assume that the choosing self in the original position is an abstract, unencumbered, pre-social subject — capable of stepping outside all particular attachments and evaluating principles from a morally neutral point of view. Sandel argues that this picture is a fiction. In reality, personal identity is constituted by the communal ties, histories, languages and values in which the individual is embedded. We do not choose our deepest commitments the way we choose consumer preferences; they shape who we are prior to choice.

It is important to see that Sandel does **not** endorse the crude communitarian claim that the community or majority will should automatically prevail over the individual, nor the cultural relativist thesis that no culture should ever be judged by external standards. Rather, Sandel's point is that rights *cannot* be justified without reference to some conception of the good. This directly challenges the Kantian/Rawlsian thesis that principles of justice can be vindicated by appealing only to the concept of right, independently of any ethical vision of the good life.

For Rawls (and Kant), the priority of the right over the good means:
(1) individual rights are so fundamental that the general welfare cannot override them; and
(2) the justification of these rights does not depend on any particular moral or religious worldview.

Sandel denies the second claim. What he calls the "communitarian critique" asserts that justice is not independent of the good. There are (at least) two ways to express this dependence:

1. **Justice is rooted in widely shared values** of a particular historically situated community.
2. **The justification of rights presupposes the moral importance of the human purposes those rights serve.**

Aristotle already expressed this thought clearly: before determining the nature of rights and the ideal constitution, we must first determine what counts as the most desirable form of life. Without such a horizon, the concept of justice remains empty. We turn now to two illuminating example to explain Sandel's point of view.

Example I: the meaning of free speech

The famous case of the proposed Nazi march in Skokie, Illinois — a community with a large number of Holocaust survivors — illustrates Sandel's point. Liberal theory maintains that the state may regulate the *time, place, and manner* of speech, but must remain neutral with respect to its *content*. Mill defended free expression on utilitarian grounds: even immoral doctrines must be allowed, because free and open discussion is the best means to truth. To censor content would be to impose some citizens' values on others and would fail to respect the dignity of the autonomous chooser — understood, again, as an abstract and unencumbered self whose identity is independent of any social role.

But for Holocaust survivors, the presence of Nazi symbols is not merely one more opinion in the marketplace of ideas. It is a direct assault on their identity, and a traumatic reactivation of historical terror. Nevertheless, liberals argue: if Skokie may forbid the Nazis, then why could the segregationist communities of the American South not forbid the civil-rights marchers of the 1950s and 1960s? If content cannot matter, these two cases cannot be distinguished.

For the strict liberal neutralist, there is indeed *no* morally relevant difference. Free speech must be protected in both cases, regardless of content.

Some communitarians, by contrast, would say that each community should simply enforce its own values: Skokie should exclude Nazis, and segregationist towns should exclude civil-rights activists. Sandel rejects this crude relativism as well.

His point is different: the difference between the two cases *lies precisely in the content of the speech*. The Nazi message defends genocide and hatred; Martin Luther King's marches defended equal citizenship and the dismantling of injustice. There is also a difference in the **moral worth of the communities** whose identities are at stake: the shared memory of Holocaust survivors is not morally equivalent to the racist solidarity of white supremacists.

On Sandel's view, the state cannot pretend to moral neutrality. It must sometimes uphold values that sustain a shared common good, and refuse to protect speech whose purpose is to destroy the moral foundations of civic life.

Free speech, therefore, is not an absolute procedural right detached from substantive moral goods. The legitimacy of free expression depends on the ethical quality of the ends it serves.

Example II: the meaning of religious liberty

Why should freedom of religion enjoy constitutional protection?
For the liberal, the answer is straightforward: religious liberty matters because liberty matters *in general*. Religious convictions, on this view, are simply one category of personal preference. What makes them worthy of respect is not their *content*, but the fact that they are freely chosen. They deserve protection because they express the autonomy of the individual — the priority of the right over any particular conception of the good.

Sandel thinks this is an impoverished account. It flattens religion into a mode of consumer choice and neglects the role that religious commitments actually play in people's lives. For many citizens, religious practice is not just an optional lifestyle preference. It is central to their identity, their moral self-understanding, and the shape of their lives as a whole. Religious duties are not chosen externally to what is good for them — they constitute that good.

This means that the case for religious liberty cannot be adequately captured by an abstract principle of neutral choice. What makes a religious practice worthy of respect is not the *process* by which it was acquired (choice, revelation, education) but the *place* it occupies within a good life, the virtues it cultivates, and the kinds of character it promotes. It is tied to the "thick moral sources" of a community.

If we assimilate religious liberty to liberty *in general*, we lose the ability to distinguish morally significant cases. For example: the right of an Orthodox Jewish soldier to wear a yarmulke in military service cannot be assimilated to a general right for troops to wear any kind of hat. Likewise, granting Native American communities access to peyote for sacramental purposes is not equivalent to granting a general right for recreational drug use. The moral equivalence breaks down.

Sandel's claim is that we cannot understand the meaning of religious liberty without reference to the substantive goods at stake. The state cannot be morally neutral here. Some practices express deeply rooted ethical traditions that form citizens and sustain civic virtue — and for that reason deserve protection. Others do not. Rights, in this domain as in others, cannot be justified independently of the good.

Implications of Sandel's view for social policy and global policy

Sandel's central thesis is that rights cannot be justified without reference to substantive conceptions of the good. If this is true, then social policy and global policy cannot simply be procedural or "neutral" frameworks for individual choice. They must, inevitably, take stands on moral purposes. This has several important consequences.

First, social policy cannot be morally neutral. Welfare policies, education policies, the construction of public law and the shaping of public culture cannot merely protect private autonomy; they will inevitably shape shared values. Educational policy, for example, cannot be indifferent between cultivating civic virtue and merely providing skills for the labour market. Family policy cannot ignore the ethical importance of caregiving, solidarity, or reciprocity. And debates over free speech cannot be decided without evaluating the moral worth of the causes being protected. Sandel's view pushes politics away from purely procedural liberalism and back toward substantive discussions about ends, purposes, and goods.

Second, public institutions must sometimes promote civic goods. If the "thick" goods of a community are the background that makes meaningful agency possible, then institutions have a role in cultivating those goods. This does not imply an authoritarian state, but it does imply that public schools are not neutral spaces, that civic participation is a public good to be nurtured, and that policy can legitimately privilege practices that form democratic character — such as public service, civic rituals, commemorations or the cultivation of collective memory. The contrast with Rawls is sharp. Rawls wants legitimacy to remain independent from any particular moral anthropology. Sandel replies that such neutrality is unsustainable.

Third, Sandel's view has implications for global policy. A global order cannot simply export a single liberal model and present it as "universal". Different political communities are constituted by different moral traditions — Confucian, Hindu, Muslim, indigenous, "Western" liberal Protestant-derived, and so on. This means that there may not be a single model of development or democracy that fits all societies. Global institutions cannot assume that neutrality equals legitimacy, and cross-cultural dialogue must engage with moral substance, not only procedural rights. In Sandel's terms, global ethics should be a forum of contested goods — not simply a technical distribution of "rights constraints".

In short: social and global policy must take moral sides. They cannot claim legitimacy while pretending to be value-neutral, because the very justification of rights presupposes a vision of the good.

Sandel's critique of liberal neutrality brings us to a third major communitarian figure: Michael Walzer. While Sandel shows that rights cannot be justified without reference to substantive moral goods, Walzer develops this insight into a theory of moral pluralism grounded in the diversity of actual communities. Walzer's central claim is that moral and political concepts have "thick" meanings rooted in particular histories, languages and shared understandings, and only a much "thinner" set of meanings can be shared across cultures. In this sense Walzer takes the communitarian critique one step further: if Sandel rejects the idea of a morally unencumbered self, Walzer rejects the idea of a single universal vocabulary of justice. For him, the moral world is not unified but plural. It is to this theory of "thick and thin" morality that we now turn.

6.4 Michael Walzer: Spheres of Justice

Michael Walzer (1935) is one of leading communitarians along with Taylor and Sandel. His most important contribution on social justice is *Spheres of Justice* (1983) where he develops the theory of complex equality.
According to this theory justice is not measured by one standard but rather requires that each good be distributed according to his social meaning, and we should not allow any good (like money or political power) allowed to dominate or distort the distribution of goods in other spheres; further, justice has meaning in a particular nations or community and cannot be developed as a generalized moral standard to be applied everywhere. In his *Thick and Thin – Moral Argument at Home and Abroad (1994)* revised and extend his previous work on justice and frames his ideas about justice, social criticism and identity in a globalized world.

Thick and Thin

In *Thick and Thin: Moral Argument at Home and Abroad* Michael Walzer develops a distinctive communitarian theory of moral pluralism. His central idea is that moral concepts are always embedded in social meanings. Every community has a "thick" morality — a dense network of shared understandings, symbols, memories, and practices — which gives content and shape to what justice and the good life mean in that particular historical society. But at the same time Walzer argues that there is a much "thinner" morality that we can share across different cultures. These thin moral terms are minimal, abstract, and largely negative (for example: do not massacre, do not torture, do not enslave). This thin morality is the basis for universal condemnation of atrocities, even when the thick moralities of societies differ profoundly.

This leads to Walzer's concept of **moral minimalism**. Minimal morality does not tell us how to live well; it only tells us how not to treat other human beings. It draws the basic boundary between what is morally unacceptable anywhere and what is subject to cultural variation. When we move from this thin minimal morality to substantive distributive questions — how goods, offices, honour, welfare or recognition ought to be distributed inside a society — we enter the domain of **maximalism**. Theories of distributive justice are always maximalist, because the principles of distribution depend on the shared meanings of goods within particular communities. To decide who should have political office, or how welfare ought to be distributed, we must interpret the thick values and social meanings which that community actually recognises. Distributive justice cannot be universal in form: it must be interpretive, local, historically situated.

Because of this, the social critic is not an external universal judge who stands "outside" and evaluates all cultures neutrally. The true social critic, for Walzer, is always someone who argues *from within* the thick moral vocabulary of his or her own society. Maximalist criticism is internal criticism. When we condemn injustice, we do so by appealing to the values that our own community already acknowledges, even if only implicitly. We "remind" the community of its own deeper commitments. Moral improvement begins inside traditions, not above them.

This distinction between thin minimalism and thick maximalism has powerful implications for international politics. Minimal morality can support international intervention against genocide or mass atrocities, because such acts violate the thin norms that all moralities share. But beyond this minimal horizon, the world is not a single moral community. Talk about global distributive justice is always at risk of becoming another form of cultural imperialism if it assumes that there is one universal thick morality. Walzer rejects the liberal hope for a cosmopolitan universal distributive scheme: justice among strangers must remain minimal and cautious. Beyond stopping horrors, the world cannot legislate a single correct way of life.

Finally, Walzer rejects the Rawlsian, Kantian idea of a purely self-choosing, disembedded moral subject. The self is not a homogenous universal chooser but a **divided self** — a subject formed through multiple memberships: family, nation, congregation, class, profession, language community, historical memory. Our moral identity is not derived from a single abstract rational faculty, but from these overlapping communities of belonging. Therefore justice must remain plural. There is not one universal moral horizon, but many thick moral worlds coexisting, intersecting, and negotiating with a thin layer of minimal shared norms holding them together at the edges.

Walzer's conclusion is not relativism — but pluralism. There are certain thin prohibitions which hold universally, and many thick moralities which remain irreducibly particular. Justice is therefore both human and local: minimal in the world, maximal at home.

Spheres of Justice

Walzer's distinction between "thick" and "thin" moralities sets the stage for his most influential systematic work, *Spheres of Justice* (1983). If moral meanings are always embedded in particular cultures, then distributive principles cannot be universal and uniform. They must reflect the specific social meanings that different goods have within different societies. *Thick and Thin* tells us that maximal justice is local; *Spheres of Justice* explains **how** that local justice works. It offers a theory of "complex equality" in which different goods belong to different "spheres", each governed by its own criteria of distribution. In this way Walzer develops a new model of distributive justice rooted not in a single universal principle, but in the plurality of social goods and shared understandings that make them meaningful.

The Social Meaning of Goods and the Pluralism of Social Goods

In *Spheres of Justice* Walzer's starting point is that every distributive question is interpretive. There is no single abstract category of "primary goods" (as in Rawls), valid everywhere and for everyone. Instead, goods have **social meanings**, and these meanings differ from society to society. Goods are not only material objects — they are the medium through which human relations are expressed and negotiated. "Goods enter our minds before they enter our hands." This means that before we decide *who* should get *what*, we must ask what the good *means* to this community. Medical care, for example, expresses a social meaning of alleviating suffering — so it should be distributed according to need. Political office expresses public trust and legitimacy — so it should be allocated through democratic procedures. Education expresses the cultivation of human capacities — so merit and civic commitment may be relevant criteria. What is crucial is that **different goods belong to different distributive spheres**, and each sphere has its own internal logic of justice. We cannot use the same distributive principle everywhere.

Because meanings are historical and cultural, justice is plural: what counts as a just distribution changes over time and between societies. There is no universal "currency" of justice, no master-principle that allows us to rank all goods under one single metric (money, utility, liberty or "primary goods"). Thus, justice is always local and interpretive. Yet Walzer is not naïve: he recognises that the spheres are constantly violated. In almost every society, some goods are used illegitimately to dominate other spheres. Wealth is used to buy political office; political power is used to secure economic advantage; religion has been used to obtain legal privilege. Walzer insists that these violations are *not* accidental — they are systematic. Throughout history, different goods have functioned as dominant ones: religion in medieval Europe, land in agrarian societies, money in market societies, technology in the modern world. But total dominance is never fully achieved — the claim to complete domination by one good is always ideological.

From this diagnosis follows Walzer's positive concept of **complex equality**: a society is just when no single good can "convert" into all others, and when each sphere protects the autonomy of its own logic. Justice is therefore not about inventing a single principle of distribution, but about preventing tyranny — the tyranny of letting one kind of good rule all the others.

Complex equality is his alternative to the liberal aim of a single egalitarian distribution of resources. In a society characterised by complex equality, no single social good is allowed to dominate all the others. Inequalities may exist *within* particular spheres — for example, some people may be more admired artists, or more skilled surgeons — but those inequalities cannot be "converted" into dominance across all dimensions of life. A great painter should not, by virtue of artistic fame, gain automatic political power; a wealthy entrepreneur should not be able to buy legal exemptions; a priest should not be entitled to economic privilege. Complex equality seeks to block those conversions.

Thus, equality is not the absence of all differences. Walzer does not deny that societies will always have rankings, honours and asymmetries internal to spheres. Instead, what he rejects is the **tyranny of dominance** — the situation where one good becomes the master key that opens every other sphere. Complex equality is the arrangement in which *the social meaning of each good determines the appropriate criteria of its distribution*, and no good is allowed to escape its own sphere and rule others. This is why distributive justice for Walzer is not primarily about redistribution of income, but about the integrity of boundaries.

Walzer therefore gives us a model of equality compatible with pluralism. There is not one idea of the good life, not one ranking of goods, not one metric that can be applied everywhere. Equality in this view means that **no one can dominate others across all domains**, because no good is allowed to dominate all others. Justice is achieved when different kinds of social goods circulate within their own moral logics — when the diversity of social meanings is preserved rather than flattened into a single scale.

Complex equality does not require a perfect levelling; it requires a *balanced distribution of powers*, where each sphere protects itself against invasion. Justice, then, is not merely a matter of how much people have, but of **what they can do with what they have** — and whether any good can be illegitimately converted into another.

Blocked Exchange

If goods have different social meanings, then not all goods can be legitimately exchanged with one another. Walzer's term for this is **blocked exchange**. The point is simple but powerful: to allow certain goods to circulate freely — to be bought, sold, converted, traded — would corrupt what they *are*. Their meaning would be destroyed by the mode of acquisition.

Money, for example, can legitimately buy consumer goods or houses or leisure. But it should not be able to buy political office, judicial verdicts, school grades, or marital love. Why? Because the social meaning of political power is one of *democratic authorization*, not private purchase. The social meaning of justice is *impartiality*, not personal profit. The social meaning of love is *affection, intimacy, commitment*, not market calculation. When money crosses those boundaries, it is an *invasion of spheres* — one good (money) tries to dominate others (power, justice, love). This is precisely the kind of dominance complex equality tries to prevent.

Blocked exchange therefore protects moral boundaries. It says: **these goods cannot be bought** — not because we dislike markets in general, but because a market in these goods would destroy the thing itself. If political power can be bought, democracy collapses into oligarchy. If legal verdicts can be purchased, law becomes merely a commodity. If families are reducible to market contracts, intimacy loses its meaning. Some things are corrupted by market logic.

So, blocked exchange is not anti-market. It is a defence of the plurality of moral meanings. It is a barrier against the imperialism of money — the danger that the most convertible good (money) claims dominance over all others. Blocked exchange is therefore one of the key institutional mechanisms by which complex equality is sustained.

The Community as a Good to Be Distributed

For Walzer, the political community is not merely the background against which distribution happens — it is itself one of the most valuable social goods. The community is the shared space of meaning within which the social meanings of other goods take shape. It is where language, shared memory, historical experience, and cultural symbols are woven into a collective identity. Although goods and people do move across national borders, the political community remains the nearest thing we have to a common world of thick meanings. That is why membership — belonging — is itself a good in need of just distribution.

To distribute the good of membership means two things: physical admission and political recognition. A passport, residence rights, citizenship — these are the institutional forms in which belonging is allocated. This is not something that can be assigned by an external authority. Membership only has meaning if the decision comes from within — from the community deciding who may enter and who may join. In this sense, immigration rules are part of the distributive logic of the community. They determine how the good of belonging circulates.

Membership — Members and Strangers

For Walzer, distributive justice always presupposes a bounded world. We cannot talk about the distribution of goods in the abstract, because distribution is always something that happens **within** a political community — among people who already share a common life. The most fundamental social good, therefore, is not money, or healthcare, or education — it is *membership* itself. Membership determines who is included in the circle of sharing, and who remains outside it.

Those who lack membership — the stateless — live in a condition of radical vulnerability. They may still enter markets, buy and sell goods, and enter private contracts, but they remain outside the network of public protections and public goods. They are excluded from the collective provision of security, welfare, public healthcare, education, and political representation. In this sense, affluent nations operate like elite universities — countless outsiders apply for entry, but only some are admitted.

Membership is not something citizens distribute to themselves (they already have it); it is something they extend — selectively — to outsiders.

Walzer also notes that when two strangers meet, moral life cannot rely on thick communal bonds. The ethical principle that remains is minimal: one should help when another is in need and when the cost or risk of helping is low. This minimal "obligation to aid" is a thin morality — the Good Samaritan principle. But a world governed only by that thin morality, without political communities, would mean that justice collapses into sheer non-coercion, private charity, and goodwill. It would be a world of global libertarianism — without shared institutions and without enforceable solidarity.

This is why communities must regulate their own populations: through immigration policy, admission criteria, and even birth control policies. These practices are not merely administrative. They are part of the way a community distributes *its most precious good* — belonging. Membership is the gate that determines who participates in the internal system of complex equality, and who remains a stranger at the door.

The only imaginable alternative to the political community would be the universal community of all humankind — *humanity as such* — but this does not yet exist as a functioning moral and political world. Cosmopolitan theorists sometimes argue for global justice based on hypothetical contracts between all persons. Walzer's reply is that such a global contract would have no real authority unless power was somehow centralized at a world level. But if one dismantles the political monopolies of states in order to enforce global rules, then global power would inevitably concentrate — not in a pluralistic network of communities — but in the hands of a single bureaucratic elite. This would not achieve complex equality, but rather **simple** equality enforced from above — stripping away local meanings and local self-determination.

Thus, for Walzer, political membership is not a trivial boundary issue. It is one of the most important goods to be distributed — because without a community of shared meanings, there is no interpretive context in which any distributive principle could make sense at all.

Countries as Clubs

Walzer pushes this logic further by suggesting that political communities can also be understood analogously to *neighbourhoods, clubs, or even families*. Like private associations, states develop principles of inclusion, shared identity, and internal cohesion. Henry Sidgwick —who briefly imagined humanity as a single world-community — already recognized in the 19th century the limits of such cosmopolitan idealism: a completely open world would undermine the patriotic sentiments that sustain civic attachment; it could generate economic dislocation by depressing wages for the most vulnerable; and it might weaken the cultivation of shared intellectual and cultural traditions. For Walzer, the community's ability to regulate its own boundaries is therefore essential to maintaining the possibility of common life. If we simply tear down national borders, we are unlikely to produce one global neighbourhood — we are more likely to generate a proliferation of *small and hostile fortresses*. Membership selection, therefore, is analogous to club admission: there are rules, criteria, and expectations — and the central normative question is always: *with whom* do we wish to share social goods? Moreover, the analogy sometimes shifts from clubs to *families*: many states prioritize kinship ties in their immigration law, treating ancestral affinity as a special claim (as in the reunification of Germany, or the return of ethnic Germans from Eastern Europe after 1945). For Walzer, then, immigration is not a private consumer choice of individuals; it is an act of collective self-determination by the community itself. Countries decide what kind of moral identity they wish to reproduce — and membership is one of the most precious goods that they distribute.

Refugees and Guest Workers

Walzer applies the general principle of membership also to contemporary debates over immigration — particularly refugees and guest workers. With respect to **refugees**, Walzer argues that our obligations do not derive merely from humanitarian sentiment in the abstract. In some cases, we owe them the same kind of solidarity we owe fellow nationals — especially when we bear responsibility for the conditions that produced their exile. The clearest example is the American reception of Vietnamese refugees after the U.S. war in Indochina. But even when the persecuted are not "our" people and do not share our political culture, a moral duty of *mutual aid* remains when the numbers are manageable. This is precisely what nineteenth-century Britain practiced: it sheltered political dissidents (including anti-liberal socialists) because the guiding principle was not ideological affinity but protection of the oppressed. Today, these obligations are codified into international conventions on asylum and non-refoulement — but for Walzer the normative logic was already implicit in thin moral duties that pre-date the treaties.

The case of **guest workers** exposes a different kind of injustice. Many advanced societies admit migrant labour on a *temporary* basis, and construct institutional systems that keep these workers economically integrated while politically excluded. They pay taxes, work in the economy, and contribute to social wealth — but remain non-participants in the political community. Walzer believes this violates the principle of political justice. The state's internal process of self-determination — the process through which a people shape the distribution of their social goods — must be open, at least in principle, to all those who live within its borders. A permanent class of non-citizen labourers is, therefore, morally indefensible. Participation in the economy should ultimately open a pathway toward participation in the polity. Guest workers must not become a permanent caste. They must, to use Walzer's phrase, be *set on the road to citizenship*.

In short: Walzer rejects both the purely humanitarian charity model (which treats admission as optional generosity) and the economic instrumental model (which treats migrants as disposable labour). Membership is a social good. If we rely on people's labour — and if we live with them in the same territory — they must eventually be invited fully into the shared life of the political community.

Cultural Relativism of Justice — Walzer contra Rawls

For Walzer, justice is not a universal formula but a *context-dependent* practice. Different societies attach different meanings to their goods, and therefore justice requires attention to those particular, historically formed meanings. To ask how education, health care, political office, land or religious status should be distributed, one must first ask what these goods *mean* within that specific community. Justice, in this sense, is always interpretive rather than purely procedural. This position stands in direct contrast to Rawls's methodological abstraction in *A Theory of Justice*, where the attempt to derive universal principles from an "original position" ignores the cultural embeddedness of goods.

Walzer's critique of Rawls is, therefore, not primarily about Rawls's conception of the person. It is a goods-based critique. Rawls's "primary goods," which supposedly serve any rational life plan, fail precisely because they presuppose that goods have trans-cultural, stable meanings. For Walzer, they do not. We are *culture-producing creatures* — we inhabit thick, meaningful worlds. Justice is rooted in the distinct understandings of offices, honours, roles, and things that make up a shared social life. To override these meanings in the name of a universal distributive formula is not neutrality — it is domination.

The implications are considerable. Walzer rejects the idea that meritocracy or market pricing can serve as universal rules of distribution. Markets may distribute consumer goods efficiently, but they cannot legitimately determine who receives medical care or political office. Transparent "spheres" must be protected: public funding for education is justified not only instrumentally, but to prevent the economic sphere from corrupting the knowledge sphere; similarly, political finance must be regulated to prevent money from converting into political power. In these ways Walzer's view supports robust welfare state policies — but not on the basis of universal rights. The case for welfare begins with *membership*: "welfare rights are fixed only when a community adopts some program of mutual provision." Redistribution is, therefore, a civic practice, not a universal entitlement.

At the same time, Walzer and Rawls share important commitments. Both believe civil and political rights should be equal, both insist that educational opportunity should not depend on birth, and both reject the concentration of wealth when it distorts political equality. But Rawls grounds these commitments in a universal liberal theory of right. Walzer grounds them in the social meanings internal to shared forms of life. Justice, for Walzer, is always plural — because the world's goods, and the world's cultures, are plural.

6.5 Conclusion — Communitarianism and the Future of Global Justice

Taken together, Taylor, Sandel and Walzer challenge the liberal hope for a *view from nowhere*. They do not deny the importance of rights, nor do they relapse into authoritarian majoritarianism. Rather, their common move is to insist that persons are never abstract choosers floating above the social world: identity is formed through language (Taylor), through moral commitments and shared purposes (Sandel), and through historically embedded social meanings of goods (Walzer). If this is true, then justice cannot be grounded in a purely procedural or content-neutral liberal framework. Communities matter — because shared moral horizons, civic practices, traditions and histories are the foundations from which individual agency becomes meaningful.

This has deep implications for debates about global justice. Communitarianism does *not* simply defend national closure. But it insists that talk of “global justice” must proceed through the interpretive engagement of actual political communities — not by imposing abstract norms from above. Taylor shows that recognition is always culturally situated; Sandel shows that public reason cannot be severed from moral substance; Walzer shows that distribution depends on the social meanings internal to each community, and that membership itself is the most basic good.

The communitarian challenge to liberal cosmopolitanism is therefore this: a just global order cannot be built by asserting universal principles while ignoring the thick moral vocabularies that constitute peoples. If global justice is to be more than a technocratic slogan, it must become a forum of *contested goods*, negotiated among communities that take their own moral self-understandings seriously. In this sense, communitarianism does not close the door to universalism; it forces universalism to become *dialogical* — rooted in practices of recognition, civic membership, and the plurality of ways in which human beings understand what is good.

Appendix A: Aristotle as Background: Teleology, Justice, and Flourishing

Aristotle provides a foundational philosophical backdrop for both contemporary communitarianism and the capability approach because he denies the possibility of separating questions of justice from questions concerning the good life. His account of justice is explicitly **teleological**, oriented toward the ends internal to practices, and **honorific**, concerned with identifying the excellences that such practices ought to recognize.

1. Teleology as the Structure of Aristotelian Justice

For Aristotle, any inquiry into what justice requires must begin with an examination of the **telos**—the purpose or defining function—of the practice or institution under consideration. Entitlements cannot be determined independently of these purposes: the mode of distribution appropriate to a good follows from an understanding of what the good is *for*. Only after the telos is specified can one determine the relevant criteria of desert.

Once the purpose of a practice is identified, Aristotle maintains that one must then determine the **virtues or excellences**that the practice is meant to honour. Distribution is governed by those traits that are constitutive of the practice's proper functioning. Thus, Stradivarius violins should be allocated to the most accomplished violinists not because doing so maximizes aggregate welfare, but because the practice of violin-playing itself adjudicates excellence in that domain. Conversely, qualities irrelevant to the practice—wealth, lineage, aesthetic attractiveness, luck—have no standing in determining just distribution.

Aristotle also insists on the differentiation of **distinct spheres of justice**, each governed by its own principles of allocation. Desert, need, and free exchange apply in different domains and cannot be substituted for one another without distorting the underlying purposes of the respective practices. Political office is not purchased; medical care is not assigned by competitive skill; market exchange does not determine access to honour. Each sphere has its own internal standards.

2. The Anti-Neutralist Commitment

This teleological orientation stands in marked contrast to the Kantian–Rawlsian commitment to **neutrality**, which aims to articulate principles of justice without recourse to substantive accounts of the good. Liberal proceduralism brackets conceptions of virtue, merit, and excellence in order to secure a fair framework for individuals to pursue their private ends. For Aristotle, however, such bracketing is conceptually incoherent. Debates about justice are necessarily debates about the purposes of institutions and, therefore, about the constituents of the good life.

This anti-neutralist stance is the genealogical source of both **communitarian criticisms of procedural liberalism** (e.g. Taylor, Sandel, Walzer) and of the **capability approach's emphasis on human flourishing** rather than preference satisfaction. Aristotelian justice always links two elements—*the good to be allocated* and *the person to whom it is allocated*—and the principle of equality is indexed to the respect relevant to the telos of the good.

3. Politics as a Formative Practice

Aristotle's teleology extends directly to the political domain. Determining who ought to rule requires first articulating the **purpose of ruling**. If political association exists merely to secure mutual protection or facilitate exchange, then efficiency might suffice as a criterion for distribution of authority. But if the polis exists to cultivate civic virtue and enable deliberation about the common good, then those best equipped for such deliberation deserve political authority.

For Aristotle, politics is inherently **formative** rather than merely instrumental. The polis exists "by nature" because human beings are creatures capable of reasoned speech, whose flourishing depends on shared deliberation about ends. Hence he rejects political arrangements that reduce the state to a wealth-maximizing mechanism (oligarchy) or a system of unmediated preference aggregation (pure democracy). Such models misconstrue the higher telos of political association: the promotion of *eudaimonia* as an active, virtuous form of life.

4. Contemporary Resonances: Communitarianism and the Capability Approach

Aristotle's framework illuminates contemporary disputes that appear procedural but are in fact **teleological**. The debate over affirmative action, for instance, hinges on competing conceptions of the **purpose of the university**. If the university exists solely to cultivate and reward scholarly excellence, admissions should be determined strictly by academic merit. If, however, the university is also charged with sustaining a democratic polity, widening access, or representing social diversity, then diversity may count as a relevant institutional excellence. The apparent conflict over fairness thus reduces to a disagreement about ends.

This logic underlies two significant modern theoretical developments:

1. **Communitarianism**, which argues that justice cannot be detached from the social meanings and purposes embedded in practices; and
2. **The capability approach**, which conceives justice as enabling individuals to develop and exercise their human capacities rather than merely protecting formal liberties or allocating resources.

Both traditions inherit Aristotle's central claim: **justice is evaluative, not neutral; it presupposes an account of human flourishing and of the excellences appropriate to shared institutions.**

5. The Enduring Legacy

Aristotle's model is teleological and honorific: justice asks what an institution is for and which excellences it ought to reward. While some of his substantive judgments (e.g. his defence of natural slavery) are untenable and collapse on his own premises, the structure of his reasoning remains influential. Misidentifying the telos of a practice leads to injustice; proper identification grounds legitimate principles of distribution.

For this reason, Aristotle should not be regarded merely as a historical figure but as the **genealogical ancestor** of two of the most prominent alternatives to Rawlsian liberalism:

- **Communitarian contextualism**, which emphasizes the embeddedness of goods in practices; and
- **The neo-Aristotelian turn to flourishing** (Nussbaum, Sen), which foregrounds the development of human capabilities as the primary metric of justice.

Both inherit and elaborate the Aristotelian thesis that **justice is inseparable from an account of human nature and from the purposes of the institutions within which human flourishing is realized.**

Chapter 7

INSTITUTIONAL COSMOPOLITANISM

7.1 Introduction

The contemporary debate on global justice has been shaped decisively by a set of theorists who reject the territorial boundaries of classical liberal thought. Among the most influential of these figures are Charles Beitz and Thomas Pogge. Both begin from the insight that the moral and material fate of individuals today is no longer determined chiefly by what happens within their own state. Global trade regimes, financial flows, investment rules, debt arrangements, resource privileges, supply chains and security alliances now structure human prospects across borders. The world, in effect, already functions through institutional patterns that are global in scope. If so, then the conditions Rawls identifies domestically as the site of justice — a structured scheme of cooperation that shapes life chances — exist internationally as well.

Yet Rawls's *The Law of Peoples* denies this implication. Rawls deliberately restricts his principles of distributive justice to the domestic sphere, and at the global level offers only a thin "duty of assistance." His model is explicitly non-distributive and statist. It treats "peoples" rather than persons as the basic units of moral consideration, it tolerates large inequalities between societies once a minimal threshold has been reached, and it refuses to recognize the global economy as a morally relevant basic structure. For Beitz and Pogge, this position is too conservative: it fails to register the empirical fact of global interdependence, and thus fails to acknowledge that the same structural features which trigger distributive duties domestically now exist across the entire global order. And, above all, it remains silent before the persistence of extreme poverty and the huge global inequality that characterized the global world.

This chapter examines how Beitz and Pogge move beyond Rawls. Beitz, in *Political Theory and International Relations*, first argued that global economic interdependence transforms the international realm into a field of distributive justice. Pogge further radicalizes this insight by developing an explicitly institutional approach: he claims that global poverty is not merely unfortunate but is systematically produced — foreseeably — by the institutional order we collectively impose. Both thinkers thus relocate the site of justice from the bounded state to the global institutional structure itself.

Institutional cosmopolitanism is therefore not simply an extension of Rawlsian justice outward. It represents a shift in the unit of moral analysis: from "peoples" to *persons*, from domestic jurisdiction to global design, and from acts of charity to obligations grounded in shared institutional responsibility. It is the faulty design of the global order with its trade rules and tariffs, patent policies, World Bank and IMF policies, pollution rules, acquisitions from rich resources countries by multinational companies at low prices, absence of labor-standard worldwide that produce massive human rights deficits condemning billions of people to severe poverty. In this way, Beitz and Pogge insist that justice must be global because our institutions already are.

7.2 Charles Beitz and the Reframing of Global Justice

Charles Beitz is one of the most influential early theorists of cosmopolitan global justice. A student of John Rawls, he is often placed — together with Thomas Pogge — among the founders of contemporary institutional cosmopolitanism. Beitz's major contribution is the translation and extension of Rawls's *A Theory of Justice* from the domestic sphere to the international one. His 1979 book *Political Theory and International Relations* is widely recognised as the starting point of philosophical global justice as a research field.

Before Beitz, the dominant tradition in international theory was Realism. Realists — following Thucydides, Hobbes, Machiavelli — argue that the fundamental drivers of global politics are conflict, power, and survival. States are the primary actors; the system is anarchic; there is no world sovereign; and therefore states rationally maximise their own security and advantage. On this view, international institutions, NGOs, and individuals matter far less than states. Normative questions about who owes what to whom are pushed aside. The classic realist tradition is therefore *silent* on international distributive justice.

Beitz breaks this silence. He begins by noting that the world has changed: we now inhabit a single, interdependent global economic structure. Famine, radical inequality, and environmental destruction are not accidental — they are shaped by global rules, practices, and institutional arrangements.

Beitz therefore rejects the idea that moral obligations stop at borders. While citizens in wealthy societies have long been told they owe "charity" or "mutual aid" to distant sufferers, Beitz argues that this is not enough. For utilitarians, there is no theoretical distinction between humanitarian assistance and justice — but Beitz wants something more foundational: distributive duties grounded in a conception of social justice.

His central move is this: the same principles of justice that govern distribution within a domestic society should apply *globally* if the international realm exhibits the same morally relevant features — namely: **interdependence**, **cooperation**, and **institutional structure**.

Beitz does not deny that Rawls's contractarianism was revolutionary domestically. But he contends that Rawls artificially restricts his contract model to the nation–state. For Beitz, this restriction is no longer coherent. The international system increasingly resembles a domestic basic structure — rules, norms, interdependence — and therefore global distributive principles become morally required.

Beitz's claim is thus bold and direct:

international distributive obligations are obligations of justice, not merely of charity.

The burdens and benefits of the global economy must be assessed and justified with reference to persons as persons — regardless of citizenship. And today's massive inequalities are not morally neutral facts: they constitute the central political challenge of our coming century.

Beitz, in short, opens the cosmopolitan turn by insisting that Rawls's principles must be taken seriously beyond the state — because the structure of the world already is.

Rawls famously conceives society as "a fair system of cooperation" for mutual advantage. Individuals depend upon shared institutions to secure important social goods — security, opportunities, income, self-respect — while also presenting competing claims to them. This is precisely why principles of justice, in Rawls's model, are needed: to regulate the distribution of benefits and burdens that arise *within* such a cooperative scheme.

For Rawls, this model is strictly bounded by the nation–state. His two principles of justice apply to the basic structure of *domestic* society — to the institutional framework that citizens share. On this view, justice is a matter internal to a bounded community. The behaviour of states in international affairs follows not from a principle of global distributive fairness, but from the domestic justice of well-ordered societies. If each society organizes itself justly at home, Rawls believes the world will remain stable and peaceful.

Cosmopolitans — and Beitz is among the first — reject this restriction. They argue that the world economy now functions as a cooperative venture for mutual advantage in ways that are structurally analogous to Rawls's domestic basic structure. Even if states were fully autarkic, Beitz notes, the moral problem would not disappear: the *natural* distribution of resources itself generates moral conflict.

In Beitz's global original position, the parties know that natural resources are not distributed evenly across the planet. Oil, fertile land, timber, minerals — these are not located according to any moral criteria. They are, to use Rawlsian vocabulary, *morally arbitrary*. This mirrors the way Rawls treats natural talent domestically. You are not more worthy because you happen to possess musical genius. Likewise, a country is not more deserving simply because it sits atop oil reserves. Since the distribution is arbitrary, the benefits reasonably flowing from such resources — like the benefits from natural talents — are candidates for redistribution under a principle of justice.

Beitz therefore proposes a **global resource redistribution principle**, functioning internationally as the difference principle functions domestically: resource advantages should work to the benefit of the globally least advantaged.

But Beitz's argument does not stop at the hypothetical case of autarkic states. The empirical world is *far from* autarkic. We inhabit a dense web of trade relations, capital flows, intellectual property regimes, travel, multinational corporate activity, and treaty-based international rules. Globalization — whatever mix of benefits and harms one attributes to it — has created vast positive-sum outputs: growth, productivity increases, technological diffusion. But these benefits are distributed unevenly. States are positioned very differently — politically, economically, structurally — within the global order. International property rights, trade rules, and debt regimes systematically shape who gains and who loses.

This is the crucial step: **international interdependence creates a pattern of benefits and burdens that would not exist if states were isolated**. This interdependence therefore has the same moral structure as Rawls's domestic "cooperative scheme." And if the moral structure is the same, then justice cannot logically be restricted to the state.

For Beitz, *global distributive justice follows from the mere fact of global socio-economic interdependence*.

Accordingly, if we re-run Rawls's original position at the global level — without assuming membership in a particular national society — the parties would choose principles that apply to all persons, not merely compatriots. They would not allow citizenship or resource endowment to determine moral worth. And if the difference principle is valid domestically because it regulates morally arbitrary factors (like talent), then consistency requires that a **global difference principle** be chosen internationally.

This yields a familiar — but radical — conclusion:

A tax transferring resources from resource-rich states to resource-poor states is not charity; it is distributive justice.

For Beitz, national borders have no independent moral weight. The structure of the world economy — not the map — determines the scope of justice.

7.3 Thomas Pogge

Thomas Pogge is another central figure in the development of institutional cosmopolitanism. Like Beitz, Pogge studied with Rawls at Harvard — but he goes further than Beitz in arguing that normative evaluation must begin not with individual conduct, but with the justice of the *basic institutions* that shape people's life chances. Institutions — not isolated individual choices — structure who lives well, who is protected, and who suffers avoidable harm. For Pogge, any credible moral theory must therefore examine how well our institutional order performs for the *least advantaged*.

Pogge's central claim is simple but radical: if Rawls's principles are compelling domestically, they must also apply internationally — because the global order is already an institutional structure that systematically shapes the opportunities, risks, and vulnerabilities of persons across borders. Global poverty is not merely an unfortunate natural fact; it is the outcome of the rules we impose.

Pogge has developed his framework across several important works, including: *Realizing Rawls* (1989), *World Poverty and Human Rights* (2002 / 2008) and *Politics as Usual* (2010)

These books together articulate Pogge's institutional approach: negative duties not to harm through unjust global rules, the idea that global cooperation generates obligations of justice, and the argument that we — especially citizens of affluent states — are often complicit in maintaining institutions that foreseeably and avoidably produce severe poverty.

In addition to these theoretical contributions, Pogge has also developed practical global reform proposals. Two of the most well-known are:

- **The Health Impact Fund** — a mechanism to reward pharmaceutical companies based on the *health impact* of their innovations, while selling approved medicines worldwide at production cost.
- **The Global Resources Dividend (GRD)** — a small global tax on the extraction or use of natural resources (beginning with petroleum), designed to redistribute a portion of resource value to the world's poorest populations.

These proposals are concrete attempts to operationalize cosmopolitan justice in the realm of global policy and institutional design. In this section of the chapter, however, we will focus on Pogge's *theoretical* contribution — his institutional conception of human rights and global justice. His practical proposals for reform — including the Health Impact Fund and the Global Resources Dividend — will be discussed later in this book in the chapter on global poverty (Chapter 8).

7.4 World Poverty and Human rights

Pogge's disagreement with Rawls is not a rejection — it is a revision. He remains deeply indebted to Rawls's framework, especially to the idea that justice is fundamentally about how institutions shape life chances. Pogge even acknowledges that Rawls's own "duty of assistance" could, in principle, be used to criticize affluent states for failing to meet the basic subsistence needs of poorer societies. But Pogge insists that this misses the central moral insight.

The problem is not merely that wealthy societies help *too little*.
The deeper injustice is that they *harm too much* — through the global rules they collectively design and enforce.

As Pogge puts it, the present international order foreseeably and avoidably condemns billions to severe poverty, and since the end of the Cold War hundreds of millions have died as a result of poverty-related causes. The core of Pogge's critique is therefore structural: affluent countries are not neutral bystanders to global misery; they are active contributors to an institutional architecture that produces it.

This is the pivot point where Pogge departs from Rawls: from charity to duty, from assistance to institutional responsibility, from domestic justice to global justice.

In thinking about how justice might apply beyond state borders, John Rawls famously makes a sharp distinction between domestic and international contexts. His domestic theory — developed in *A Theory of Justice* — is built around moral individualism: the basic unit of moral concern is the individual person. Yet in his later work on international relations, Rawls shifts. Thomas Pogge has argued that this shift is theoretically damaging and ultimately undermines the universal reach of justice.

Pogge begins with Rawls's use of the "original position" at the international level. Rawls does not simply extend the domestic device. Instead, he redesigns it. In the international original position, the deliberating parties represent "peoples," not individual persons. This marks a retreat from the very moral individualism that Rawls previously insisted upon. Moreover, Rawls never clarifies what counts as a "people": is it a state's population, or an ethno-cultural nation that may not map onto state borders? In addition, Rawls restricts who is allowed to participate at all: only "well-ordered" societies — liberal or "decent" — are admitted as equal members of the Society of Peoples. "Outlaw states," "burdened societies," and other non-liberal orders are excluded in advance. In doing so, Rawls denies equal standing to precisely the societies that would most need just terms of cooperation.

Rawls also weakens the veil of ignorance in his international thought experiment. Whereas in the domestic case parties do not know their social position, status, or conception of the good, in the international case representatives are still permitted to know whether they represent a liberal or a decent society. Finally, Rawls instructs these deliberators not to design global institutions of distributive justice, but merely to agree on rules for good conduct in foreign policy. Rawls does not attempt a theory of global distributive justice at all; rather, he simply seeks to outline the foreign policy ideals of a reasonably just liberal people.

Rawls's defence for this limitation is straightforward: there is no global "basic structure" analogous to the institutional framework of a domestic society, and any attempt to construct one — especially in the form of a world-state — would risk tyranny or collapse into civil strife. Rawls invokes Kant here, claiming that Kant too rejected a world-state.

Pogge replies that this reading of Kant is historically weak. Kant rejected empire by conquest, not a voluntary federal union of republics. More importantly, Pogge argues that even if a world-state were empirically unlikely or even undesirable, this does not justify withholding principles of justice at the global level. Rawls's own criterion of social justice is explicitly meant to *assess* and *critique* institutional arrangements — not to presuppose a specific institutional structure in advance. Justice should govern institutions, not wait until the right institutions happen to exist.

The world today is already tied together by relations of deep interdependence: trade regimes, economic regulations, financial flows, diplomatic bargaining, supply chains.

These cross-border structures generate benefits, costs, and vulnerabilities that profoundly shape people's life chances. On Pogge's view — and on similar arguments by Charles Beitz and Kok-Chor Tan — such dense interconnection is sufficient to trigger distributive duties globally. If global institutions can determine whether some populations flourish while others remain trapped in deprivation, then justice must speak here.

Pogge therefore believes that Rawls's refusal to globalize his own principles is not merely an analytic error — it also introduces a normative inconsistency. Domestically, Rawls does not tolerate illiberal doctrines or comprehensive views that deny persons equal citizenship. But internationally, Rawls insists that liberal peoples must "tolerate" decent hierarchical societies — even though those societies deny their own members basic liberal freedoms. This means Rawls ends up protecting the autonomy of collectivities rather than the freedom of individuals. For Pogge, this results not in principled pluralism, but in a double standard that sacrifices the liberties and opportunities of countless human beings for the sake of accommodating non-liberal political cultures. It also legitimates a world in which extreme global inequalities are left unexamined — or worse, made invisible by Rawls's theoretical architecture.

Pogge's critique thus goes to the heart of the emerging debate that this book traces: whether justice can remain credible if it remains territorially confined. Rawls's international theory, Pogge argues, narrows the moral horizon of justice at precisely the historical moment when global structures most decisively shape human life.

Pogge's Moral Cosmopolitanism and the Scope of Global Responsibility

Thomas Pogge is one of the most important contemporary defenders of *moral cosmopolitanism*. The core intuition behind cosmopolitanism is simple yet radical: every human person affected by an institutional arrangement deserves equal moral consideration. The status of the person — not culture, not citizenship, not nationality — is what ultimately grounds moral obligation.

Pogge identifies three features essential to any cosmopolitan view.
First, **individualism**: individuals, and not groups, are the basic units of moral concern. We may care about families, nations, or religious communities — but only because they are composed of persons. Second, **universality**: equal moral regard applies to all persons without exception. No category of exclusion — whether based on race, class, gender, ethnicity, or citizenship — can be morally legitimate. Third, **generality**: the equal moral status of persons has global reach. Moral concern does not stop at national borders.

From this follows an important conclusion: global boundaries are morally arbitrary. Citizenship has no more normative weight than skin colour or social class. It cannot determine who deserves priority in the distribution of rights, resources, opportunities, or protections. For Pogge, moral assessment must be based on principles that apply to all persons equally — and not on the accident of where they were born.

It is this cosmopolitan starting point that grounds Pogge's critique of John Rawls. Whereas Rawls's domestic theory assigns ultimate moral standing to individuals, Rawls's international theory shifts the focus from *persons* to *peoples*. In the "international original position," Rawls no longer treats individuals as the primary bearers of claims, but instead models justice around the equality of collectivities. For Pogge, this is a fundamental retreat: if individuals are the basic units of moral concern, then the baseline for distributive justice must be the condition of the *worst-off persons in the world*, not merely the worst-off within each state. Otherwise, vast global inequalities remain unchallenged.

To navigate between what he calls "weak" and "strong" cosmopolitanism, Pogge develops a distinctive middle position — *intermediate cosmopolitanism*. Strong cosmopolitanism requires radical global equality; weak cosmopolitanism only insists on minimal threshold duties, such as Rawls's limited "duty of assistance." Pogge rejects both. He wants to preserve the reality of special relationships (one may indeed owe more to one's own children than to strangers), but without allowing these ties to weaken universal obligations. His solution is to distinguish between **negative duties** and **positive duties**.

Negative duties — such as the duty not to harm — apply equally to compatriots and foreigners. They are stringent, universal, and non-negotiable. Positive duties — such as assisting others in need — may legitimately vary in strength depending on one's relationships and social ties. One may have stronger positive reasons to help one's compatriots first — but this cannot erase one's negative duty not to impose or uphold unjust institutions that harm others abroad. Pogge's claim is powerful: affluent societies today do not merely fail to help the poor — they are often complicit through their global institutional arrangements in *harming* them.

This is why Pogge expands the institutional approach to justice beyond state borders. For Rawls, institutions are primarily domestic. For Pogge, today's global economic order — from the WTO and IMF to multinational corporations and trade regimes — already constitutes a de facto structure with enormous distributive effects. Avoidable global poverty, in this context, is not an unfortunate accident; it is the foreseeable outcome of institutional design.

Nevertheless, Pogge's institutional approach faces its own challenges. Authors like Simon Caney (2005) argues that if duties of justice apply only where there is institutional membership or causal harm, then individuals outside these structures — for example, victims of natural resource scarcity, disease, or domestic tyranny — may fall outside the scope of justice altogether. Caney therefore insists that duties of justice also extend beyond institutional relationships and may generate positive duties that are not conditional on causal complicity.

Despite these debates, Pogge's contribution is decisive. He places the global institutional order — and the harms it produces — at the centre of contemporary moral philosophy. Moral cosmopolitanism, on Pogge's view, forces us to recognise that our obligations do not stop at national borders. Our institutions, and our participation in them, make us responsible for the world that they help to create.

Pogge on Human Rights and the Moral Responsibility to End Global Poverty

This chapter now turns to Thomas Pogge's moral account of global poverty. Pogge's central claim is striking: today's severe poverty is not simply the result of local misfortune, corruption, or poor governance, but is largely *produced* and *maintained* by unjust global institutions. If global structures contribute to the persistence of deprivation, then global institutions must be reformed so that the poor are not systematically denied access to their basic needs.

For Pogge, extreme poverty is avoidable and foreseeable — which means it is morally imputable. Alternative institutional arrangements are possible, and there exist versions of the global order in which the poor would not be harmed to the same extent. His proposal therefore is not merely charitable aid, but institutional redesign: a transformation of global rules and practices such that human rights are actually realized. Article 28 of the Universal Declaration of Human Rights is central here: "everyone is entitled to a social and international order in which the rights and freedoms set forth in this Declaration can be fully realized."

Pogge interprets this as a normative mandate. The fulfilment of human rights is not the full content of global justice — but a necessary core. Institutional schemes are unjust if they fail to secure the realization of human rights *where this is reasonably possible*. In other words: human rights form the universal minimum standard of basic justice.

In developing this argument, Pogge is not trying to enumerate or legislate a complete list of specific human rights. Nor is he concerned with metaphysical debates about the ontological "existence" of rights. What matters for him is the *moral notion* of human rights — the idea that there are certain fundamental claims persons hold *simply as persons*, claims which any legitimate political order must respect.

Human rights are not created by governments. Rather, governments have legitimacy only when they protect pre-existing moral human rights. The language of the UDHR itself presupposes this: legal codification does not create rights, but recognizes them.

Pogge traces this notion historically. Human rights evolved out of the older language of natural law and natural rights. All three frameworks express weighty moral demands that apply to persons everywhere, across cultures and epochs. Natural law emphasized conformity to a divinely ordered cosmos; natural rights shifted focus toward the protection of persons as moral claimants; human rights secularize this idea further and mark a movement from metaphysical grounding toward a more explicitly political vocabulary.

Despite this shift, Pogge maintains that human rights remain weighty, universal moral requirements. They apply equally to all human beings, and violations matter equally regardless of whose rights are being violated.

Some critics worry that this move from "natural rights" to "human rights" risks making rights seem more fragile — as if what is humanly conferred might also be humanly revoked. Pogge's response is that human rights are "weighty, unrestricted, and broadly sharable moral concerns." They have normative force not because of metaphysical origins, but because they matter deeply to human beings across diverse cultures, traditions, and institutions.

The authority of human rights is grounded not in their metaphysical origin, but in their universal normative significance.

In this way, Pogge reframes discourse on global poverty: not as a humanitarian appeal to generosity, but as a matter of justice. Fulfilling human rights is the minimum threshold for an institutional order to be morally acceptable — and when our global institutional order predictably fails to do this, then we are not merely passive observers of global poverty. We are implicated in it.

Pogge's Institutional Conception of Human Rights

For Pogge, human rights are best understood not as direct claims against individual moral agents, but primarily as claims on *coercive social institutions* — and secondarily as claims on those who uphold and maintain these institutions. This marks a crucial distinction between two ways of thinking about human rights. On the **interactional** approach, human rights are claims directed toward individuals or collective agents whose behaviour must be morally assessed. On the **institutional** approach, human rights are standards by which we assess the design and operation of social structures. Human rights, in this view, are principles for evaluating the basic rules that structure social cooperation.

In institutional cosmopolitanism, individuals bear responsibility not because they personally perpetrate harm, but because they help design, sustain, or benefit from unjust social orders. The duty is therefore indirect: people must not cooperate in imposing or maintaining institutional schemes that foreseeably and avoidably leave human rights unfulfilled.

Domestic institutions therefore have the primary role in fulfilling human rights. Yet in a globalized world, institutional responsibility cannot stop at national borders. Today's dense web of global interdependence — trade regimes, financial institutions, multinational corporations — means that the actual conditions under which human beings live cannot be traced solely to local politics or isolated governments. The realization of human rights is shaped by global structures — and so the corresponding duties fall partly upon them, and upon those who influence them.

To say that someone has a right to *X* means, for Pogge, that any coercive institutional order affecting them must, so far as reasonably possible, be designed so that all persons have *secure access* to *X*. Human rights are therefore violated officially when institutions fail to protect secure access to these goods above a certain reasonable threshold. Individuals who cooperate in sustaining those institutions share responsibility for that official disrespect.

Pogge defines harm in a stringent way. The global poor are harmed when their basic human rights remain unfulfilled; when this deprivation is causally connected to institutional design; when those who helped create or maintain those institutions could foresee the harm; when alternative institutional arrangements exist that would avoid comparable deprivation; and when such alternatives are known or knowable. In such cases, the obligation to help is compensatory, grounded in responsibility for institutional harm.

This approach helps explain why Pogge thinks human rights violations are fundamentally *official* violations. Threats or occasional abuses, by themselves, do not constitute official disrespect if the institutions in place adequately secure access to the relevant goods. A single torture case, for instance, does not necessarily prove a systemic denial of the right to physical security; a legal system that routinely leaves its population exposed to arbitrary violence does.

Critics — such as Wenar (2005)— argue that Pogge's reliance on probabilistic security may generate counter-intuitive outcomes. A regime could, for example, provide secure access to most citizens while officially authorizing atrocities against a few individuals. Pogge's view, critics say, must be refined so that certain direct official acts — such as torture — always constitute rights violations regardless of broader statistical security.

Even so, Pogge maintains that human rights need not be legally codified to be fulfilled. A society could secure access to nutrition, education, or basic liberties through a range of legal and non-legal arrangements — social norms, cultural solidarities, dispersed ownership, or welfare provisions — without formal rights legislation. Conversely, legal codification does not guarantee fulfilment if institutions fail in practice.

Pogge does not deny the relevance of interactional duties. Individuals still owe negative duties not to harm others, and positive duties to support institutional reforms where needed. But the institutional view avoids both libertarian minimalism (which denies duties where one is not directly harming) and utilitarian maximalism (which would make everyone responsible for every deprivation anywhere).

Pogge's human rights framework is deliberately thin. Human rights protect the conditions necessary for persons to live minimally worthwhile lives while allowing different societies to expand and specify additional rights according to their own values. Key rights include physical integrity, political participation, subsistence, basic education, freedom of movement, and opportunities for economic participation.

Human rights, therefore, are fundamental standards for evaluating the legitimacy of power. A state that fails to secure these rights for its population loses moral authority; its power becomes mere force rather than rightful rule.

In summary, Pogge's conception of human rights expresses a commitment to oppose institutional arrangements that fail to respect the basic needs of persons. Because global institutional orders shape life chances everywhere, the responsibility to secure human rights is likewise global.

Freedom from Poverty as a Human Right

Pogge argues that freedom from severe poverty should be understood as a *human right*. In doing so, he refers specifically to those living in conditions of extreme deprivation — those for whom even a minimally adequate diet and basic non-food essentials are unaffordable. This category also includes those who live only marginally above that minimal threshold but whose lives remain continually threatened by insecurity and deprivation.

This claim is controversial. Libertarians maintain that our only universal moral duty is the negative duty not to harm others. If rights require corresponding duties, libertarians argue, then a right to basic necessities seems implausible: it is unclear who could possibly be held responsible for providing food, shelter, or medicine to all human beings. On this view, failing to assist the poor — even if morally regrettable — is not itself a rights violation.

Utilitarians take the opposite stance. If one can help without sacrificing something comparably important, one ought to do so. But critics reply that such a principle is excessively demanding, and neglects the fact that human moral life is structured through partial relationships — family, community, and nation — which legitimately shape our patterns of concern and obligation.

Pogge responds by rejecting the assumption that a right must necessarily specify *who* must supply the object of that right. Rights are ultimately rights to certain forms of conduct — actions and omissions — and can therefore be specified in more than one way. If one particular specification of a right entails implausible duties, that does not invalidate *every* possible formulation of that right.

For Pogge, the question becomes clearer when we adopt the institutional view. The duty that corresponds to the right to basic necessities is not the duty of individual affluent persons to directly feed or clothe the poor. Rather, it is the duty not to cooperate in sustaining an unjust global order that foreseeably produces avoidable misery. Since today's global institutional rules predictably contribute to massive and preventable poverty, those who participate in these institutions — affluent states, corporations, and citizens — have a negative duty not to impose such harm. This negative duty grounds a positive responsibility: to reform global institutions so that they no longer deprive the poor of their basic human rights.

Thus, on Pogge's view, freedom from poverty counts as a human right not because the rich must individually supply the needs of the poor, but because no one may participate in global arrangements that foreseeably perpetuate extreme deprivation. The duty is structural — not charitable — and it is grounded in the shared design and maintenance of the global order itself.

7.5 Conclusion: The Promise and Limits of Institutional Cosmopolitanism

Beitz and Pogge transform the debate on global justice by insisting that we take seriously what our world already looks like: a densely interdependent web of institutions whose rules shape life chances across borders. If Rawls taught us to see the domestic basic structure as the primary subject of justice, institutional cosmopolitanism extends this insight outward. The global economy, on this view, is not a morally neutral backdrop but a basic structure in its own right – one that distributes benefits and burdens, creates vulnerabilities, and helps determine who enjoys secure access to their human rights and who does not. Justice, accordingly, cannot stop at the border because the institutions that most decisively affect people's lives do not.

Charles Beitz opens this path by arguing that once the world economy functions as a system of cooperation for mutual advantage, the Rawlsian logic that applies within states must also apply between them. Natural resource endowments and positions in the global division of labour are as morally arbitrary as natural talents within a society; they therefore cannot be allowed to determine, unchecked, who lives in affluence and who remains in permanent deprivation. Beitz's global original position and proposal for a resource redistribution principle symbolise this shift: what earlier appeared as international "charity" is reconceived as a matter of distributive justice, owed to persons as such rather than to "peoples" as collective units.

Thomas Pogge radicalises this move by reframing global poverty not as a failure to help, but as the outcome of structural harm. His institutional cosmopolitanism centres on negative duties: affluent states and their citizens have stringent obligations not to impose or uphold global rules that foreseeably and avoidably perpetuate severe poverty and human rights deficits. On this account, the "normal" operations of the WTO, IMF, debt regimes, resource privileges, and intellectual property rules are not morally innocent. They are key mechanisms through which avoidable deprivations are produced and reproduced. Human rights, interpreted institutionally, thus become standards for judging the legitimacy of the global order itself: an international scheme that could be reasonably redesigned to avoid massive poverty, yet is not, stands condemned as unjust.

This shift from interactional to institutional responsibility has profound implications. It relocates moral focus from isolated acts of charity to the design of trade regimes, financial rules, and resource governance; from the generosity of individuals to the justice of the background frameworks they sustain. In doing so, institutional cosmopolitanism exposes the limitations of both Rawls's Law of Peoples, with its statist "duty of assistance," and libertarian views that recognize only negative duties between individuals but ignore the moral significance of the global rules they collectively endorse. By foregrounding the official, structural character of human rights fulfilment and violation, Beitz and Pogge give us a language to describe global poverty, not as tragic misfortune, but as a systemic injustice for which agents can be held responsible.

Yet the chapter has also highlighted important tensions and open questions. By tying duties of justice so closely to institutional membership and causal involvement, institutional cosmopolitanism risks neglecting those who suffer from deprivations that are not easily traced to global rules – for example, victims of domestic tyranny, environmental catastrophe, or internal conflict. Critics such as Simon Caney (2005) worry that a framework built only on institutional complicity may leave some urgent needs outside the scope of justice and reduce positive duties to a secondary, optional status. There are also questions of feasibility and legitimacy: global reforms like a Global Resources Dividend or a Health Impact Fund require powerful states and corporations to accept constraints on their advantages. How such measures can be secured democratically, and how they can respect pluralism and collective self-determination while pursuing cosmopolitan aims, remains contested.

At the same time, institutional cosmopolitanism can seem normatively demanding and psychologically distant. Ordinary citizens may find it difficult to see themselves as responsible for abstract global structures, and causal pathways from their everyday choices to distant harms are often opaque. The challenge, then, is not only to diagnose structural injustice, but to articulate forms of political agency – through social movements, democratic contestation, consumer action, and transnational advocacy – that can connect individual motivation to institutional transformation. Without such pathways, cosmopolitan theory risks becoming a powerful moral critique with limited grip on political practice.

Despite these difficulties, Beitz's and Pogge's contributions mark a decisive turning point in the philosophy of the global world. They compel us to recognise that our obligations to others are not exhausted by episodic aid or humanitarian sentiment; they are embedded in the ongoing design and maintenance of the rules under which we all live. Institutional cosmopolitanism thus reframes global justice as a question of how to construct and reform the basic structures of the international order so that the human rights of all persons can be secured. The following chapters will build on this insight, applying it to specific domains such as global poverty, climate change, and migration, and asking how a cosmopolitan ethics of institutions can be made concrete: not only as a critical standard for judging the present, but as a practical guide for imagining more just global arrangements.

PART III

Chapter 8

WORLD POVERTY

8.1 Introduction

"About 8.5 percent of the global population lives in extreme poverty in 2024. This means that 692 million people worldwide live on less than $2.15 per person per day. While the extreme poverty rate fell from 38 percent in 1990 to 8.5 percent in 2024, it has stalled more recently amid lower economic growth and multiple shocks such as the COVID-19 pandemic, high inflation, and increased conflict and fragility." (The Poverty, Prosperity and Planet Report 2024, World Bank Group).

If we look beyond the narrow $2.15 line, the picture becomes even more troubling:
• at $3.65/day → ~1.7 billion people (21.4% of humanity) are poor in 2024
• at $6.85/day → ~3.5 billion people (43.6%) live in poverty — almost half the world

Despite declining *percentages* at higher thresholds, the *absolute number* of people under $6.85/day has barely shifted since the 1990s because population growth offset improvements.

Forecasts make this trajectory even more bleak. By 2030, extreme poverty is projected to still be 7.3% — more than twice the World Bank target of 3% — and only 69 million people are expected to escape extreme poverty between now and then. Under current growth patterns, the 3% goal is out of reach for decades; even optimistic growth scenarios do not get there before the 2040s.

Extreme poverty has devastating and wide-ranging consequences for human life. The most striking indicator is the huge gap in life expectancy between rich and poor countries. Today, a child born in a high-income country such as Japan or Switzerland can expect to live around 82–85 years. By contrast, in many low-income countries — such as Sierra Leone, Chad, or Central African Republic — life expectancy is still around 52–55 years. This difference of almost thirty years is not an abstraction: it translates into millions of premature deaths every single year.

The mechanisms behind this gap are structural and cumulative. People in poor countries are chronically undernourished: according to recent UN estimates, around 735 million people in the world are undernourished in 2024. Millions lack access to essential medicines — the WHO reports that nearly two billion people have no regular access to even basic pharmaceuticals. Safe drinking water remains unavailable to roughly two billion people, and about 3.5 billion still lack safely managed sanitation.

Illiteracy, especially among women and girls, remains pervasive: UNESCO estimates that more than 760 million adults cannot read or write. Children in low-income settings are far more likely to be engaged in wage labour outside their household — around 160 million children are currently in child labour according to the latest ILO figures. These are not marginal disadvantages — they constitute an immense, continuous, and incommensurable form of suffering.

Given these realities, almost no one disputes the moral claim that "we should do something about" global poverty. There is broad agreement that extreme world poverty constitutes a grave injustice that must be corrected. What is contested, however, is not the urgency, but the diagnosis: scholars, economists, and philosophers disagree deeply about the causes of persistent poverty and about what remedies are most justifiable, most effective, and most legitimate.

In this chapter we begin by analysing one central mechanism for addressing global poverty: Official Development Assistance (ODA). ODA institutionalises the idea of "a duty to assist" across borders — it is, in principle, the official channel through which richer states support poorer states. In practice, however, ODA is often shaped by national interest, strategic competition, and geopolitical priorities, which has generated intense debate about whether — and to what extent — aid truly serves the poor.

After discussing the political economy of ODA, we then turn to alternative moral frameworks for responding to global poverty. In particular, we examine (1) the consequentialist utilitarian approach associated with Peter Singer , and (2) the rights-based approach associated with Thomas Pogge. These two perspectives illustrate the broader tension between "doing good outcomes" (maximising welfare) and "not violating rights" (not contributing to, or benefiting from, unjust global structures).

8.2 The Scale of Global "Help": ODA, Remittances, and Philanthropy

In order to understand the role of aid in world poverty and inequality, we need to situate official development assistance within the wider landscape of international financial flows. In 2024, total Official Development Assistance (ODA) from OECD-DAC donors amounted to **USD 212.1 billion**. The United States remained the largest donor in absolute volume (about USD 63.3 billion), followed by Germany (USD 32.4 billion), the United Kingdom (USD 18.0 billion), Japan (USD 16.8 billion) and France (USD 15.4 billion). If we correct for economic size, the Scandinavian donors are the most generous: Norway (1.02% GNI), Luxembourg (1.00%), Sweden (0.79%) and Denmark (0.71%) all exceeded the long-standing 0.7% target. Yet even within these numbers the composition matters: in-donor refugee costs and humanitarian spending absorb substantial shares, and not all of this money reaches development programs in recipient countries.

What is less widely appreciated is that ODA is not the largest external resource for poor households. **Remittances** — the money that migrants send home to their families — were estimated at around **USD 685 billion** to low- and middle-income countries in 2024. These flows go directly to households and have strong poverty-reduction effects.

In this sense, remittances are both larger and more directly poverty-relevant than most aid, and they are counter-cyclical: when domestic economies deteriorate in the global South, migrants often send more money.

A third category is **private philanthropy and NGO funding**. Here the data are harder to compare, because there is no unified global reporting system. One market estimate valued the global NGO/charity sector at roughly **USD 329 billion** in 2020 — but this figure must be used cautiously: it includes domestic giving in rich countries, hospitals, universities, churches, cultural institutions, animal welfare groups and many other organisations that have nothing to do with development or poverty. A better indicator for "philanthropy for development" is provided by the OECD Centre on Philanthropy. Their most comprehensive mapping found that large international foundations channelled roughly **USD 42.5 billion between 2016 and 2019** — about **USD 10–11 billion per year** — to developing countries, with **health** by far the largest category (around 43% of that total), followed by education and agriculture. If we take this breakdown as a guide, roughly sixty per cent of philanthropic development spending can be considered to fall into poverty-relevant sectors, though even here not all funding is targeted specifically at the extreme poor.

Taken together, these numbers illustrate a picture that is more complex than the conventional story of "rich countries giving aid to poor countries". ODA is large, but remittances are larger still. Private philanthropy is meaningful, but far smaller, and unevenly distributed by sector. Only a fraction of NGO or foundation activity is focused on poverty. Understanding this landscape is crucial if we want to evaluate the real scale — and the real limits — of external finance in addressing global inequality.

8.3 A Critical Literature Review: Does Foreign Aid Reduce Poverty?

The academic literature on the developmental effects of foreign aid is vast, but its findings are not harmonious. What emerges from recent studies is not a simple yes-or-no answer, but a picture of deep **heterogeneity**: aid does different things in different institutional settings. In terms of scale, Official Development Assistance (ODA) is substantial — around USD 200–220 billion annually in recent years — and in some countries it has represented the majority of public expenditure. Yet the core analytical question remains: *what does aid actually do* for growth and poverty reduction?

Older macro-econometric work in the 1970s and 1980s tended to ask whether aid was correlated with growth. Results were famously inconclusive: some studies found small positive effects, some negative, most statistically insignificant. The "micro–macro paradox" (Mosley 1987) crystallised the puzzle: many individual projects scored well by donor evaluation criteria (clinics built, wells drilled), but national growth performance remained poor. Later cross-country work — Burnside & Dollar, Collier & Dollar — argued that aid worked **only** in countries with "good policies" and macroeconomic stability, but these claims did not survive replication well, and the policy-quality approach faded.

Recent empirical literature has taken two newer directions.

First, there is the **institutional turn**: many studies now argue that the key effect of aid is not on capital accumulation directly but on the incentive structure of governments. In fragile, low-accountability environments, large aid flows can weaken state–citizen fiscal links, entrench rent-seeking, and reduce taxation effort. Deaton and Moyo articulate this view normatively, and contemporary political-economy empirics reinforce it. A 2023 UNU-WIDER review of aid in fragile states underscored weak governance as the binding constraint: in such contexts, aid often has little or negative effect on growth. Recent "hump-shaped" findings (Bethencourt 2024) support a threshold view: aid may have positive effects where institutions are already moderately strong, but becomes counterproductive when institutions are weak.

Second, there is the **sectoral turn**: research disaggregates by type of aid. The evidence is strongest in **health**: targeted vertical programs (vaccines, malaria treatment, HIV therapy) have saved millions of lives. Here the literature is clear: aid has had first-order welfare effects. But this is precisely where the institutional argument re-enters. Life-saving "verticals" succeed because they *bypass* weak national systems. The challenge is not delivering penicillin — it is building tax systems, payrolls, drug procurement chains and regulatory capacity.

As a result, contemporary reviewers increasingly differentiate between **aid as humanitarian delivery** (which often works) and **aid as state-building** (which often does not). Poverty, from this perspective, is not simply a shortage of resources but a problem of political economy.

This literature also shows that **aid is not the largest poverty-relevant flow**. Remittances to low- and middle-income countries, around USD 650–700 billion per year, exceed aid several times over and are strongly pro-poor because they reach households directly. Philanthropic flows to developing countries are modest by comparison (~USD 10–11 billion per year in recent OECD mapping), and strongly concentrated in a few sectors (health, education, agriculture).

The present state of research thus supports a nuanced conclusion:

Aid can save lives, and has done so most clearly in health. Aid has had far weaker effects on long-run growth and poverty reduction when it substitutes for domestic institutions rather than supporting them. In settings with weak governance, aid may unintentionally entrench the very obstacles that hold development back.

This is not the argument that aid "never works", nor that rich countries owe the poor nothing. It is the claim — increasingly supported by recent evidence — that **the effectiveness of aid depends more on politics than on volume**. Structural change cannot simply be delivered; it must be built.

8.4 Some contrasting views

Not all thinkers regard foreign aid as a force for development. While some, like Jeffrey Sachs, see large-scale assistance as the key to ending poverty, a different strand of the debate is sharply sceptical. In this perspective, the problem is not that wealthy countries have been insufficiently generous, but that the very architecture of aid has been structurally damaging. Two of the most influential contemporary critics — Dambisa Moyo and Angus Deaton — argue that long-term government-to-government aid does not build development. On the contrary, they claim that it weakens the institutional foundations on which genuine growth and accountability depend. For them, poverty persists not because the rich have given too little, but because aid has often undermined the political and fiscal relationships that make self-sustaining progress possible.

Dambisa Moyo - "Dead Aid" and the critique of "aid as development"

One influential contemporary critique of development orthodoxy comes from the Zambian economist Dambisa Moyo. In *Dead Aid* (2009) her core claim is stark: **official development assistance – i.e. long-term government-to-government aid flows to Africa – has not generated development; instead it has actively *blocked* it**.

Her argument is not that humanitarian emergency relief is morally impermissible. She is not arguing against food drops in a famine or medical support in an epidemic. The target is a *systemic development model* in which low-income countries are placed on a permanent IV-drip of concessional aid, budget support and World Bank/IMF packages, decade after decade, as if that in itself were a path to prosperity.

Moyo argues that this model has four destructive effects:

1. **Aid entrenches dependency**
   Governments learn to govern *for donors*, not citizens. When tax revenue is not needed to finance the state, democratic accountability withers.
2. **Aid corrodes institutions**
   Large, opaque flows of politically mediated capital create rents, not productive investment. State capacity becomes oriented around extracting and distributing external inflows, not building fiscal systems, regulatory quality, or governance.
3. **Aid displaces domestic and foreign investment**
   Cheap, non-performance-linked money crowds out market finance. Why build bond market credibility if the budget arrives anyway?
4. **Aid sustains kleptocratic elites**
   Corruption is not a pathology that emerges *despite* aid but one that is structurally incentivised *by* aid.

The famous Moyo line is that *aid has actually kept Africa poor*.

For her, the relevant counterfactual is not a romanticised "autarkic Africa" — it is a different mode of integration into the global political economy: **finance via Eurobonds,**

trade, foreign direct investment and risk-bearing private capital rather than via the political and bureaucratic mechanisms of western donors.

This is not a merely empirical claim. It is also a normative one about agency.

The deeper philosophical message is: *a development model that keeps governments answerable to donors rather than citizens is not just ineffective — it is unfree.*

In the context of our broader discussion of poverty and inequality, Moyo's intervention highlights a provocative reversal: global poverty is not always a simple problem of *insufficient help*. It can also be a problem of **the wrong kind of help**, structured in ways that undermine the very institutional preconditions of emancipation.

If *Dead Aid* is a critique of how aid has worked, it is equally a proposal for how development *should* be financed. Moyo's programme is not sentimental — it is explicitly financial. The core idea is that **poverty is overcome not by more transfers, but by creating the institutional conditions in which capital must be earned, not gifted.**

She proposes several concrete shifts:

1. **Use bond markets rather than donors**
   Governments should finance long-term infrastructure the same way middle-income countries have done: by issuing sovereign bonds in international capital markets.
   This introduces a disciplining mechanism — if a government governs poorly, the price of borrowing rises. In other words: accountability is built into the cost of capital.
2. **Prefer foreign direct investment over aid**
   Capital that is invested — rather than given — carries risk for the investor and requires the creation of real assets.
   Moyo's paradigmatic case is Chinese investment in African infrastructure: not out of benevolence, but as mutually beneficial exchange. This is meant to shift relations from charity to contract.
3. **Make trade, not charity, the engine of income growth**
   Trade is not simply the movement of goods. It forces specialisation, capability development, standards upgrading and the creation of export-oriented production.
   For Moyo, openness to markets — and the removal of protectionism in the West — is a necessary global condition for development.
4. **Expand microfinance and financial inclusion**
   If the problem is a lack of productive capital at the level of small firms, the relevant intervention is not consumption-smoothing transfers but access to credit.
   Small and medium enterprise growth is treated as a structural driver of employment and innovation.

5. **Rebuild state legitimacy through taxation, not donors**
   A government that raises its revenue primarily from citizens must answer to citizens.
   A government that lives off donors answers to donors.

Hence the political dimension of her economic proposal:

shifting from aid to market finance is not only about efficiency — it is about the location of accountability.

Finally: Moyo does not call for an overnight rupture. She argues for a **time-limited phase-out**: a 5–10 year taper in which reliance on aid is reduced while private and public capital market instruments are scaled up.

In the wider philosophical frame of this chapter, the significance is this:
Moyo rejects the idea that "more helping" necessarily produces more justice.
Her claim is that certain *forms* of help — when institutionalised as permanent, non-reciprocal transfer — can entrench the very unfreedom they seek to remedy.

Whether one accepts her empirical priors or not, her intervention forces a conceptual question: is poverty sometimes not a lack of resources, but a lack of institutions in which resources matter?

Angus Deaton: aid, power, and the political economy of poverty

In recent years, one of the most influential and unsettling critiques of foreign aid has come from the economist Angus Deaton, Nobel prize in Economics. In *The Great Escape – Health, Wealth and the Origins of Inequality*, Deaton examines why some societies have broken out of poverty while others have not, and he argues that foreign aid — in the form in which it is usually delivered — is not the solution most people imagine it to be. His argument is not emotional or ideological; it is grounded in data, political economy, and the institutional realities of how aid actually moves. For Deaton, the problem is not the generosity of the rich, but the structure of aid itself: aid flows to governments, not to people, and in doing so it alters incentives and power relations inside poor countries in ways that often undermine, rather than support, long-run development.

Angus Deaton's argument about aid begins with an institutional clarification that most popular discussions ignore: **foreign aid is not a person-to-person transfer — it is a government-to-government flow**. This matters because it means aid is not simply money: it is a resource that enters the political system at the level of the executive. It strengthens the hand of those already in power, changes incentives inside the state, and affects the relationship between governments and citizens.

Against the familiar view that global poverty could be solved if the rich simply "gave more," Deaton calls this belief the **aid illusion**. The illusion rests on a hydraulic metaphor: poverty is imagined as a "shortfall" that can be filled with dollars. For Deaton, poverty is not numerical — it is **institutional**. The decisive variable for human development is not how much capital exists in a country, but how capital is organised, who controls it, and to whom rulers feel accountable.

If governments finance themselves through taxation, they must bargain with their citizens. That bargaining is the foundation of representation and of the long-run construction of state capacity. But when governments receive large external transfers, they often **avoid taxing citizens**, and so accountability flows outward to donors, not inward to voters. In this respect, aid can "crowd out" the political processes that historically built strong states.

Deaton's empirical interpretation is grounded in Africa's post-colonial economic record. Many African countries have received very high levels of aid for decades. If aid were a reliable driver of development, these should be among the fastest-growing economies on earth. They are not. Instead, growth in Africa has largely followed **commodity price cycles** — oil in Nigeria and Angola, diamonds in Botswana, uranium in Niger, gold and diamonds in South Africa, coffee in Kenya. When world prices rise, growth rises; when prices fall, growth collapses. Aid plays little role in this macro pattern — except sometimes to stabilise rulers during downturns.

This is why Deaton insists that **aid does not function like investment**. Investors require feedback, performance, returns. Aid does not. It can be given even when institutions are deteriorating, and it can insulate elites from the consequences of bad policy. In the worst cases, it strengthens the very political pathologies that keep people poor.

At the same time, Deaton is not a blanket "anti-aid absolutist." In one domain he makes a clear exception: **health**. Targeted "vertical" health programmes — vaccines, antibiotics, HIV treatments — have saved millions of lives. They work because they are technically simple, easily measured, and do not require a strong domestic bureaucracy to function. But this success does *not* generalise to the construction of hospitals, functioning public health systems, tax bases, or trained workforces. Preventing malaria is not the same as building a state.

For Deaton, the conclusion is therefore philosophical as much as economic. If development is fundamentally the emergence of accountable institutions, then **development must grow out of domestic politics** — not be injected as capital from outside. The task for wealthy countries is not to hit new numeric "targets" for giving (0.7% or 1% of GDP), but to **stop doing harm**: to stop sustaining governments that rule without their citizens, to stop treating development as a matter of plumbing, and to recognise that the deepest obstacles to prosperity are institutional, not financial. What is needed, he argues, is a better understanding of how poor societies actually work — of their institutions, their local incentives, and the complex political structures that shape their economies. Development, if it is to be real and durable, must come from within. The most useful contribution outsiders can make is often to **stop doing harm**, to stop propping up predatory governments, and to allow domestic reformers the political space to build accountable institutions of their own.

Aid that stabilises existing rulers — even when justified as humanitarian — ultimately impoverishes ordinary citizens by insulating elites from domestic accountability. For this reason Deaton rejects the idea that more aid, or different "targets" for aid volumes, are the moral solution.

Moreover, aid is not the only way rich countries impede development. Other policies in the rich world actively make poor countries worse off. Pharmaceutical patent regimes — for example under TRIPS — restrict access to affordable medicines. Agricultural protectionism in Europe and the United States distorts markets and blocks poor farmers from exporting into lucrative markets. There are, in other words, structural injustices embedded in the global economic system that produce poverty independently of aid.

It is here that Deaton highlights alternative proposals such as Thomas Pogge's "Health Impact Fund", which would reward pharmaceutical firms based on measurable health improvements rather than exclusive monopoly pricing. In exchange, firms would commit to selling their products at far lower prices. Such mechanisms align incentives with public health, rather than with scarcity and exclusion.

The underlying message is clear: **the most powerful way to help the world's poor is not to send more money, but to stop doing the things that keep them poor**.

Jeffey Sachs

Among contemporary thinkers on global poverty, Jeffrey Sachs stands almost at the opposite end of the spectrum from Dambisa Moyo and Angus Deaton. If Moyo and Deaton are sceptics of large-scale development aid, Sachs is its most articulate defender. His basic claim — developed in *The End of Poverty* and in his work on the UN Millennium Development Goals — is that **extreme poverty is solvable, and that the decisive instrument for solving it is more aid, not less**.

Sachs argues that many very poor societies are trapped in what he calls a *poverty trap*: incomes are so low that people cannot save; because they cannot save they cannot invest in the basics of productivity — fertilisers, irrigation, health, education, roads — and therefore they cannot grow. Poverty becomes self-replicating. The only way out, he claims, is a "big push" of external resources — a massive injection of capital, science, medicines and infrastructure — to lift societies over the threshold beyond which self-sustaining growth becomes possible. For Sachs, then, development is not primarily a political problem of domestic accountability, as it is for Deaton and Moyo, but a technical engineering problem: arranging inputs in the right sequence and at the right scale.

This is where he sharply contrasts with Moyo. For Moyo, aid creates dependency, corrodes institutions and entrenches unaccountable elites. Sachs believes the failure of aid programmes is not that there is *too much* aid, but that there has never been *enough* — certainly not enough to break the poverty trap. Where Moyo sees aid as the poison, Sachs sees insufficient aid as the missed cure. (What Deaton fears is simply that the entire macro-model is wrong: the rope by which we try to pull poor countries out of the pit may be the rope that strangles their institutions.)

Sachs also differs from Deaton in his trust in planners and experts. Deaton insists development cannot be "delivered" from outside because the key variable is political accountability — something that cannot be purchased or externally imposed. Sachs, by contrast, speaks in the language of diagnostics, scaling and targeted intervention. In this sense he treats development like a solvable medical problem rather than a deeply political one.

Thus we have two distinct diagnoses:

- For Sachs, **poverty is primarily a lack of capital**, and aid is the missing input.
- For Deaton and Moyo, **poverty is primarily a problem of institutions**, and aid undermines those very institutions.

In this debate, Sachs stands for the "optimistic interventionist" position: that extreme poverty is not a tragic mystery, but a solvable engineering problem. What divides him from Moyo and Deaton is not compassion, but political economy. The question therefore remains open: is global poverty a shortage of money — or a shortage of the institutional conditions that make money productive?

Reforming ODA: three competing pathways

Given the range of criticisms that have emerged about aid — from the charge that it fosters dependency, to the claim that it props up corrupt elites, to Pogge's deeper argument that aid distracts us from institutional harm — an obvious question surfaces: should Official Development Assistance simply be abandoned? Interestingly, the contemporary policy debate gives a clear answer: the vast majority of scholars and practitioners do not want to abolish ODA, with the exceptions of Moyo and Deaton's reservations. The real disagreement concerns **how it should be reformed**. What we now see is not a rejection of aid, but rather a contest between competing visions of what aid ought to become.

One dominant reform trend is represented by the OECD DAC agenda and the Paris–Accra–Busan "aid effectiveness" framework. Here the problem is not aid per se, but *bad* aid — fragmented, donor-driven, opaque. The solution is therefore technical, institutional, and managerial. Aid should be aligned with recipient national strategies, coordinated among donors, transparent, and rigorously evaluated. This line is closest to the development economics of Banerjee & Duflo: the problem is not whether to help — but how to help in ways that work. On this reading, good governance and evidence-based programmes can redeem ODA. The key is to do fewer things badly, and more things well.

A second vision is associated with Jeffrey Sachs. Sachs agrees that aid can be better organised — but he argues that the core problem of the last 40 years is not mismanagement; it is **under-investment**. For him, aid saves lives when it is used to fund bed nets, vaccines, antibiotics, soil supplements, teacher training, or basic infrastructure. We did not run a failed experiment in aid — we never funded the experiment properly at all. Therefore, Sachs calls for predictable, long-term financing streams, potentially backed by small global taxes (on airline tickets, carbon, financial transactions). ODA should not be marginal "charity" — it should be a structural financing mechanism for global public goods and SDG-type interventions.

A third approach, closer to Pogge, accepts that aid has a role — but insists that the primary driver of global poverty is not a lack of generosity; it is **structural injustice**. Aid will always be insufficient if the very rules of trade, intellectual property, taxation and resource sales systematically disadvantage the poor. In this perspective, ODA makes moral sense only if it is conceived not as benevolence, but as *compensation* — a partial remedy for harms that affluent states help to create. Pogge therefore does not want to abolish ODA, but he refuses to treat it as the main solution. Institutional reform must come first. ODA, in the meantime, is merely what we owe to those we have already harmed.

In short, the contemporary reform debate no longer centres on whether aid is "dead" — but on what aid is *for*. Aid can be seen as a technical instrument (OECD style), as a scaling mechanism for life-saving interventions (Sachs), or as a transitional compensation tool within a larger project of justice (Pogge). In this way, the dispute about ODA maps directly onto the deeper normative divisions that structure this chapter: Singer optimises individual giving, Sachs optimises state spending, and Pogge seeks to transform the rules that produce deprivation in the first place. What all three share — despite their theoretical differences — is the recognition that the status quo is not morally acceptable. Reform is not optional. ODA, however conceived, must evolve if it is to serve as part of a credible response to world poverty.

8.5 Peter Singer's Utilitarian Approach to Global Poverty: from States to Affluent People

In contemporary ethical debates on global poverty, two influential normative frameworks dominate the discussion: Peter Singer's consequentialist utilitarianism and Thomas Pogge's rights-based institutional critique. Although they belong to different philosophical traditions and diverge on the nature of duties — Singer emphasising our positive obligation to reduce suffering, Pogge stressing our negative duty not to sustain unjust structures — they converge on one fundamental point: the status quo is morally indefensible.

Peter Singer is perhaps the best-known contemporary philosopher arguing that affluent individuals — not only states — have a stringent moral obligation to act against extreme poverty. His position is deeply rooted in utilitarianism: if we can prevent something bad from happening, without sacrificing anything of comparable moral importance, then we ought to do it. This argument was first formulated in his seminal 1972 essay *Famine, Affluence and Morality*. There, Singer invites us to consider a simple thought experiment: if we see a child drowning in a shallow pond, we are morally required to intervene — even if it means ruining our clothes or being late for work. For Singer, the global poor are, morally speaking, "the child in the pond." Geographic distance and national boundaries make no difference. If we can prevent death from poverty-related causes at relatively low cost, failing to do so is morally wrong.

Singer later developed this claim further in *The Life You Can Save* (2009), where he gives the argument a more practical orientation. Here he identifies specific, highly effective charities capable of saving or improving lives for a few dollars per person. For Singer, what matters is not the sentiment of compassion but the outcome of action: we should direct resources to those interventions that yield the greatest possible reduction in suffering per unit of money. In *The Most Good You Can Do* (2015), this line of thought becomes fully institutionalised into a movement — Effective Altruism (EA). EA is Singer's applied utilitarianism: individuals and institutions should allocate time, money, and talent so as to maximise the total amount of good they can produce in the world. Morality, in this sense, becomes an optimisation problem.

Singer's approach differs sharply not only from rights-based theories such as Pogge's, but also from Rawlsian liberalism and communitarianism. Singer criticises Rawls for treating distributive justice as something that largely stops at national borders. Rawls recognises a limited duty of assistance between peoples, but for Singer this is far too weak: if lives can be saved at low cost, then rich societies are morally required to do much more than merely help "burdened societies" become well-ordered. Likewise, Singer criticises communitarians such as Michael Sandel for focusing on civic identity, shared values, and community membership rather than on concrete outcomes. For Singer, it is ethically irrelevant whether the global poor stand "inside" or "outside" our political community: suffering counts the same wherever it occurs. A moral theory that focuses on collective belonging without specifying what we must do to actually reduce global suffering — in Singer's view — lacks serious moral content.

Singer's utilitarian perspective thus contrasts sharply with Pogge's rights-based framework. Singer does not primarily claim that the global poor have a right that is being violated. His critique does not rest on duties not to harm. Instead, his claim is that we have positive duties to help, because the marginal utility generated by transferring resources to the poor is extremely high. For Pogge, global poverty is primarily a violation of negative duties: we are causing harm through unjust global rules. For Singer, even if all injustice disappeared, we would still have a duty to reduce suffering wherever possible. For Singer, the source of obligation is consequences — not rights.

Yet despite these deep philosophical differences, Singer and Pogge converge in empirical judgement: wealthy societies and international institutions — including the

WTO — fall far short of what morality demands. Singer criticises global trade rules not because they violate rights per se, but because they fail to reduce suffering and in many cases exacerbate it (for example through agricultural protectionism that blocks poor farmers from earning export income). For Singer, reforming these institutions is obligatory because it is one of the most efficient ways of saving lives.

In short, Singer's argument evolves from a philosophical claim (we ought to prevent suffering) to a practical methodology (measure cost-effectiveness) to a social movement (Effective Altruism). His central message remains constant: global poverty is not merely regrettable — it is morally unacceptable because we have it in our power to prevent it at relatively trivial cost.

Effective Altruism, Singer's legacy, organisational structure, financial flows, and its main critiques

The contemporary Effective Altruism (EA) movement can be seen as the institutionalisation of Peter Singer's utilitarian ethics. EA takes Singer's central idea — that we have strong moral obligations to prevent avoidable suffering when we can do so at low cost — and operationalises it through systematic cost-effectiveness analysis. The movement translates Peter Singer's arguments into practical guidance: individuals and institutions should allocate their resources (money, time, careers) toward those interventions that produce the greatest expected reduction in suffering per dollar or per hour.

EA is not one single organisation, but rather a decentralised network of specialised institutions. The Centre for Effective Altruism (CEA) functions as the community-coordination hub; GiveWell produces charity evaluations based on empirical and peer-reviewed impact research; 80,000 Hours advises on career choices; Open Philanthropy is the main large foundation funding EA-style projects. Although the movement initially focused on global poverty and global health (bed nets, deworming, cash transfers), it has expanded into animal welfare, biosecurity, and — more controversially — longtermism (existential risk reduction, especially AI safety and pandemic risk). Financially, EA now commands substantial philanthropic resources: Open Philanthropy alone has distributed over USD 3.3 billion since 2015, and GiveWell has moved more than USD 1 billion to top-recommended charities in the last decade. Private donors associated with EA (such as Dustin Moskovitz and Cari Tuna) have pledged many more billions to be disbursed in the coming years. Therefore EA is not only a philosophical movement: it is a major resource allocator within the global charity landscape.

However, this scale has generated important criticisms. Iason Gabriel (2016) argues that EA is not a single doctrine but a family of possible moral theories — some utilitarian, some not. The problem, he claims, is that EA often implicitly assumes the strongest form of utilitarian impartial maximisation, even though there are more moderate variants that would still count as "effective" while allowing space for rights, identity, and personal commitments. Amia Srinivasan (2015) goes further, criticising EA for accepting the neoliberal status quo: EA tells the wealthy how to do better charity, but does not interrogate the structural causes of injustice that make philanthropy necessary in the first place. In this view EA risks becoming a moral "palliative care unit" for a system that it never seriously challenges. And critics such as Amia (Srinivasan) and others argue that EA places too much faith in quantitative metrics — as if all suffering can be reliably reduced to spreadsheets — leaving out forms of justice that cannot be measured in dollar-per-DALY calculations.

In particolar, Iason Gabriel argues that although EA invites people to "do the most good they can," its strong utilitarian orientation leads to several moral blind spots. Because EA prioritises maximising aggregate welfare, other moral values — such as equality, rights, and fairness — tend to be treated only instrumentally, not as ends in themselves. As Gabriel notes, leading figures in the movement (such as Holden Karnofsky) openly state that "justice as an end in itself" is not a primary concern. This stance generates two risks. First, EA can end up neglecting the most vulnerable, because resources are directed toward the interventions that deliver the greatest total welfare rather than those who are worst-off. Second, EA can overlook the importance of rights and structural injustice: for example, Singer and MacAskill have defended sweatshop labour on the grounds that it increases income, while ignoring that workers also have basic rights to safe conditions — and that exploitative labour markets are themselves embedded in unjust global structures.

Gabriel therefore argues that EA should not tie itself exclusively to one moral theory. Instead of doubling down on maximising utilitarianism, the movement should adopt a more pluralistic ethical framework — cooperating with actors who care about reducing suffering *and* about fairness and rights. Some organisations, such as the Gates Foundation, already do this by prioritising the hardest-to-reach populations. If EA embraces such pluralism, it could become not only a tool for efficient philanthropy, but also a force for addressing the deeper injustices that generate poverty in the first place.

Taken together, these critiques point toward a common worry: EA inherits Singer's strength — moral urgency grounded in impartial concern — but also inherits Singer's weakness — the tendency to focus on the efficient *amelioration* of suffering rather than on the political transformation of the systems that generate suffering. In other words, EA tells us how to save more lives per dollar, but is less articulate about how to change the world so that people do

Concluding assessment of Singer's approach

Singer's contribution has been transformative. Few contemporary philosophers have made the moral urgency of global poverty so vivid. His core insight — that affluent individuals can prevent suffering and death at relatively trivial personal cost — has reshaped the ethics of everyday life, and the Effective Altruism movement has institutionalised this insight into measurable real-world action, channeling significant private funds into highly cost-effective interventions.

Yet the strengths of Singer's approach reveal its limitations. First, Singer places an extremely high moral demand on individuals. His framework is built around personal moral duties rather than collective institutional duties. But individual philanthropy remains modest relative to global need. Private philanthropy and NGO funding to global-poverty causes currently amount to around USD 10–11 billion per year — a fraction of what states allocate through Official Development Assistance (ODA), which is closer to USD 200 billion annually. In Singer's model, states are largely left in the background, and justice is pursued through voluntary individual donations. But voluntary action creates free-rider problems: unless everyone acts, the system relies disproportionately on a small minority of moral saints.

Second, Singer's strict utilitarian maximisation leaves little space for values that cannot be easily quantified — equality, rights, dignity, or the structural causes of poverty. As Gabriel, Srinivasan and others have noted, Singer gives us powerful tools to *ameliorate* suffering, but fewer tools to *transform* the institutions that generate it. EA can therefore risk becoming highly efficient symptom-management rather than structural reform. If justice requires institutional change rather than optimised philanthropy, then collective political action — not just better giving — is necessary.

In this sense, Singer's approach is normatively urgent but institutionally thin. His model shows what individuals can and should do **now**, but it does not fully theorise what states, corporations, or international institutions are obligated to do. To move from "helping" to "not harming," and from charity to justice, we require a framework that centres negative duties, global rules, and institutional responsibility. It is precisely this shift — from individual beneficence to structural injustice — that we now examine in the rights-based approach of Thomas Pogge.

8.6 Pogge's normative argument: Justice, human rights, and global responsibility

Pogge begins from a very modest normative starting point. A minimal condition of justice, he argues, is that any institutional order must not produce *massive and avoidable human-rights deficits*. This claim is entirely consistent with Article 28 of the Universal Declaration of Human Rights (1948): *"Everyone is entitled to a social and international order in which the rights and freedoms set forth in this Declaration can be fully realised."* If an institutional structure foreseeably and avoidably causes large-scale human-rights failures, then that structure is unjust.

Pogge notes that the existing UN human-rights framework places special priority on the duties of states towards their own citizens. Under this approach, governments have three kinds of duties:

1. **negative duties** — duties *not* to violate rights
   a) interactionally (e.g. do not torture)
   b) institutionally (e.g. train police to respect rights)
2. **positive duties to protect** — to protect citizens against rights-violating actors (crime, violence, private coercion)
3. **positive duties to fulfil** — by providing or facilitating access to food, healthcare, disaster relief, etc., with "progressive realisation" where resources are limited.

This "statist" model therefore assigns primary moral responsibility inward — governments owe duties principally to *their own* populations, and international duties are secondary, residual, or merely humanitarian.

Pogge's cosmopolitan alternative rejects this moral boundary.

For him, duties do not change when the person affected lives outside one's borders: foreign and compatriot alike fall under the same moral scope. The structure of duties is formally identical — negative, protective, and fulfilment duties — but with a crucial difference: the **negative institutional duty** (to not impose harmful rules) now applies *not only nationally, but also supranationally*. In other words, we have duties not to support international economic arrangements that predictably harm outsiders.

This is the heart of Pogge's view: our most stringent duty is a negative duty — a duty not to contribute to unjust global structures that foreseeably cause severe deprivations for others, wherever they live. His approach is therefore minimalistic: he is not demanding that affluent states *fully* provide for all the needs of the poor; he is requiring only that we stop participating in institutional schemes that harm them.

For the rest — the duties to protect and to fulfil — Pogge accepts the standard UN framework. What distinguishes his position is the cosmopolitan extension of negative duties to the global level. On this view, the current global order fails a basic test of minimal justice because it predictably and avoidably produces large-scale human-rights deficits. Therefore those who uphold and benefit from this global order bear moral responsibility for reforming it. Pogge then formulates his first premise:

(Premise # 1) under the present rules of the global economy, the human rights of most people remain unfulfilled.

This is not a rhetorical claim but an empirical one. To support it, Pogge cites extensive global data on the living conditions of the world's poor. For convenience, here we report the data presented at the beginning of the chapter.

Severe deprivations are widespread and persistent: according to recent UN estimates, around 735 million people are chronically undernourished in 2024; the WHO reports that nearly two billion people lack regular access to essential medicines; approximately two billion lack safe drinking water, and 3.5 billion people still do not have access to safely managed sanitation. UNESCO estimates that more than 760 million adults remain illiterate, disproportionately women and girls. Moreover, the ILO notes that around 160 million children are in child labour, often working under conditions that directly violate their basic rights.

For Pogge, these figures are not merely unfortunate background conditions. They demonstrate that the existing global institutional order systematically fails to secure even the most elementary human rights for large portions of humanity. The scale, depth, and persistence of these deprivations show, he argues, that the present global economic rules are not compatible with minimal justice. In other words: global poverty, on this view, is not simply misfortune — it is evidence that the global order is failing its most basic human-rights test.

Pogge then introduces his second premise:

(Premise # 2) it is foreseeable that most of this human-rights deficit would be reasonably avoidable with a different design of global rules.

This claim is again supported by empirical comparison. The extreme concentration of wealth in the current global economy — for example, recent estimates the *World Inequality Report 2026* show that fewer than about 60,000 ultra-wealthy individuals (the top 0.001% of the global population) possess as much wealth as the entire bottom half of the world's — illustrates how dramatically unequal the current distribution is, and how radically different it could be under alternative institutional arrangements. For Pogge, this is not simply a political slogan: it is evidence that today's poverty is not necessary. With another feasible, morally defensible global economic design, much of this suffering could be eliminated.

Pogge addresses the common counter-argument that conditions of the world poor have in fact improved — that extreme deprivation and global inequality have decreased over the last decades. He does not deny that some indicators have moved in the right direction. But he insists that this is morally irrelevant to the core question. Even if absolute poverty has declined, **the poor have not benefited proportionally from global economic growth**, and far greater progress would have occurred under fairer global rules. The decisive normative question is therefore not whether the poor are better off than thirty years ago, but **whether today's poverty is avoidable and whether we could eliminate it through better supranational institutional design**. If the answer is yes — and Pogge believes it clearly is — then we are collectively responsible for continuing to impose rules that foreseeably perpetuate massive human-rights deficits.

Pogge then addresses a second counter-argument which is commonly advocated : that differences in poverty trajectories across developing countries are primarily explained by *domestic* factors — history, culture, geography, or internal institutions. It is true, he notes, that poverty trends are heterogeneous (compare, for example, China's rapid poverty reduction with stagnation in Senegal). But for Pogge this observation does not undermine his thesis. Rather, it invites a further question: **why do severe deprivations persist systematically in certain contexts, particularly resource-rich poor countries?**

Here Pogge introduces one of his most important causal claims: the persistence of poverty is not simply due to internal weakness, but to the *collusion of interests* between international actors (such as multinational corporations) and the ruling elites of poor countries. Because of the "resource privilege" embedded in international law, whoever exercises effective control over a territory — even a junta or a dictator — is recognised as entitled to sell that country's natural resources on the global market. As a result, oil, minerals, and valuable raw materials can be exchanged for cash without any obligation that the benefits reach the local population. This global rule creates powerful incentives for coups, authoritarianism, and predatory governance. It is therefore no surprise that some of the poorest countries in the world are also among the most resource-rich. Under a different rule — for example, that resources could only be purchased from governments that meet minimal standards of democratic legitimacy — these structural incentives would vanish.

Thus, for Pogge, what looks like "domestic failure" is frequently the predictable result of international institutional design. The global rules are not neutral; they are enabling the very actors who perpetuate deprivation.

This is exactly the hinge of his argument: the injustice is not simply a failure to *assist* — it is a positive pattern of harm produced by global rules.

Pogge then argues that the present supranational institutional architecture is *causally* implicated in maintaining poverty — not merely failing to alleviate it. In his view, the global order is not neutral: it is designed in a way that systematically advantages the wealthy and predictably harms the global poor. Protectionist measures in high-income countries are a clear example: agricultural tariffs and subsidies in the EU, US, and Japan artificially depress world prices and prevent farmers in low-income countries from accessing lucrative export markets. Similarly, under the TRIPS agreement, twenty-year pharmaceutical patents — enforced worldwide through the WTO — ensure that life-saving medicines remain unaffordable for the poor. In environmental terms, poor countries are paying the highest costs of climate change despite having contributed the least to greenhouse gas emissions. Financially, the global system is structured to enable illicit financial flows: international tax loopholes and regulatory arbitrage allow multinational corporations to shift profits out of poor countries, resulting in capital flight that, according to various estimates, produces revenue losses for developing states of roughly USD 160 billion per year (Tax Justice Network – 2023/2024 "State of Tax Justice") — a figure comparable to total foreign aid. The same logic applies to natural resource privileges: resource-rich poor countries sell oil, minerals, or timber at low prices to transnational corporations, and the benefits accrue to ruling elites rather than to the population at large. Finally, Pogge notes the striking asymmetry in what is institutionalised and what is not. We have globally enforceable property rights (patent protections) but no globally enforceable basic labour standards; the international order actively protects the profits of inventors and corporations, but does not comparably protect the basic rights of workers.

From the two premises, Pogge then draws his conclusion:

(Conclusion) the present rules of the world economy systematically produce massive human rights deficits, and are therefore gravely unjust

If global poverty is severe (#1Premise) and avoidable (#2 Premise), then the global order that foreseeably produces and reproduces it is not merely unfortunate – it is morally wrongful.

This immediately raises the central question of attribution: *who is responsible* for the ongoing harm?

Pogge's answer is clear and uncompromising. Responsibility lies primarily with the governments of the more powerful states – the actors that have together shaped, negotiated and imposed the current global institutional system (WTO, IMF, OECD, the G20, etc.). These states operate in a context of intense *regulatory capture*: supranational decision-making is strongly influenced by multinational corporations, big banks, financial industry lobbies and other concentrated interest groups.

Pogge stresses that **the supranational level is particularly vulnerable to capture**:

- there is almost no democratic counterweight
- there is very little transparency
- moral restraints are easily dismissed by appealing to the idea that "international politics is a jungle" where one must defend one's competitive position

In this environment, the most powerful actors have both the strongest incentives and the best opportunities to "change the rules" in their favour. The result is a global order that secures privileges for the already advantaged and systematically harms the worse-off.

Under these conditions, Pogge argues, the governments of powerful states are **violating their most fundamental negative duty**: the duty not to impose a coercive institutional order that foreseeably produces life-threatening human rights deficits. And ordinary citizens of these states – insofar as they support, tolerate, or fail to oppose such policies – become *co-responsible* for that harm.

According to Pogge, this responsibility generates *two urgent obligations* for citizens in affluent countries:

A) we should try to influence supranational institutional arrangements so that they impose less harm on the global poor (institutional reform)
B) we should try – at the very least – to *compensate* for the harms our institutions currently impose (for example, by supporting institutional reform movements, and/or by shielding some of the poor from those harms through direct transfers)

These are not acts of charity. They are compensation and reform owed as a matter of justice.

In Pogge's view, global poverty is not simply a local tragedy. It is a predictable outcome of a global order that powerful states continue to design, maintain, and benefit from. Therefore, responsibility is shared – above all by those who hold political and institutional power, but also by citizens who permit this order to continue without adequate efforts at reform

8.7 Pogge's Institutional Proposals: Reforming the Rules that Create Poverty

Pogge's normative point is not only theoretical. He insists that the global order can be *redesigned* in relatively modest ways so that it would no longer systematically harm the poor. His central idea is that extreme deprivation is not primarily caused by a lack of generosity, but by the fact that wealthy states — through global economic rules and legal conventions — provide the structural incentives that fuel oppression and exploitation in resource-rich developing countries.

Two of Pogge's most concrete institutional proposals focus on removing what he calls the **international resource privilege** and the **borrowing privilege**. Today, whoever controls the coercive apparatus of a country — the army, the police, the capital city — is internationally recognised as the "government" of that country. This means that even regimes that come to power by coup, civil war, or foreign interference can legally sell that country's oil, minerals, or timber, and they can borrow money in its name. Pogge argues that this global legal convention creates a massive incentive for violent power seizures in countries rich in natural resources: the prize for capturing the presidency is not ideological — it is commercial. Whoever wins the palace wins the oil.

To counter this, Pogge proposes two institutional devices.

First, resource-rich developing countries should adopt **constitutional amendments** stating that only democratic governments — that is, governments which are elected and accountable — may authorise the sale or transfer of natural resources. This would give citizens the legal right to challenge contracts signed by dictators or non-representative elites. If a junta sells the country's oil, the contract would not be internationally enforceable, and future democratic governments would not be legally bound by deals made by illegitimate rulers. This removes the long-term profit value of looting.

Second, Pogge proposes the creation of an **international Democracy Panel** — an independent body of respected jurists, political scientists, and democracy experts who would not be tied to any single national interest. The task of this panel would be to monitor and judge whether a government is sufficiently representative to authorise resource sales. If the panel determines that a regime is not democratic — e.g. because of military rule, repression, or fraudulent elections — then foreign corporations and banks would be prohibited from entering into legal contracts for natural resources or for sovereign loans with that regime. This would close the global market for stolen assets. The idea is that multinationals and international banks would stop dealing with tyrants not because they suddenly "became moral," but because the legal system would make it unprofitable.

In short, Pogge wants to shift incentives "upstream." Instead of trying to fix poverty *after* it is produced — through aid and charity — he wants to redesign the rules under which global markets operate so that massive harm is not produced in the first place. His model is not charity-based but justice-based: wealthy states have duties not only to help, but above all **to stop participating in a global scheme that rewards the**

violent seizure of power in the Global South. If the resource privilege and borrowing privilege were removed, Pogge argues, the root drivers of coups, authoritarianism, and resource theft would be significantly weakened, and millions of people would gain the chance to live in societies where their natural wealth benefits them rather than the elites who oppress them.

Whether this is politically feasible remains contested, but the ambition of Pogge's proposals lies in their moral clarity. He asks us to imagine global poverty not as a humanitarian tragedy but as a predictable outcome of rules that we collectively sustain — and therefore rules we can collectively change.

These proposals — the constitutional limitation on resource sales and the creation of an international Democracy Panel — are examples of Pogge's strategy of *rule-redesign* rather than charity. They aim to remove the structural incentives that currently reward authoritarian rule in resource-rich countries. Yet Pogge's institutional imagination does not stop here. The same logic — that the global order must be adjusted so that the poor stop being structurally disadvantaged — motivates two further proposals which broaden the scope from natural resources to the global economy as a whole: the **Global Resource Dividend**, and the **Health Impact Fund**.

These two initiatives operate at a different level, but pursue the same idea: relatively small modifications to the global rules of ownership, taxation, and intellectual property could generate major reductions in avoidable suffering. Where the resource-privilege proposals focus on who is allowed to *sell* a country's wealth, the Global Resource Dividend concerns *who should benefit* from humanity's use of natural resources at all; and the Health Impact Fund aims to change the incentive structure of the pharmaceutical industry so that life-saving medicines become accessible to those who need them, not only to those who can pay high monopoly prices.

We can now turn to these two further proposals — which illustrate Pogge's belief that global poverty is not primarily a failure of charity, but a failure of institutional design.

The Global Resource Dividend

Pogge's most widely discussed institutional proposal is the **Global Resource Dividend** (GRD). This idea develops directly from one of his core philosophical claims: that humanity shares a common moral claim to the planet's natural resources. No individual, no corporation, and no state "created" the oil under the ground, the copper in the mountains, the fish in the oceans, or the atmosphere's absorptive capacity for carbon. These things are not the product of labour or merit — they are morally arbitrary starting points. Therefore, according to Pogge, it is not morally legitimate that a small minority of humanity claims exclusive ownership over resources simply because of the accident of where they happened to be born.

The GRD is meant to institutionalise this moral intuition into a concrete redistributive mechanism. Under the GRD, states would continue to exercise sovereignty over their natural resources — nothing in the proposal abolishes national ownership — but a small surcharge (a "dividend") would be levied whenever natural resources are extracted or sold on international markets. This could apply, for example, to every ton of CO_2 emitted, every barrel of oil, every ton of copper, every ton of coal, or every cubic metre of natural gas. In other words: the GRD is a very small tax on the use of the earth.

The funds collected through this dividend would be pooled into a global redistribution mechanism and transferred directly to poverty reduction programs in the poorest regions of the world. Pogge stresses that the dividend rate could be very low — perhaps 1% or even less — and yet the revenue would still be enormous, given the scale of global resource use. Even a modest levy on fossil fuels alone could generate tens of billions of dollars annually — more than current global humanitarian aid flows.

The point of the GRD is not charity. Pogge is explicit on this. The dividend is owed because the exclusive appropriation of the world's natural resources by wealthy states is institutionally coercive and morally arbitrary. A redistribution mechanism is therefore needed to compensate those whose starting position in the global lottery of birth denies them access to what should be shared goods. The GRD thus embodies Pogge's principle that, before talking about charity, we must stop *institutional harm*. The present global order enables wealthy countries to extract rents from global resources that morally belong to humankind as a whole; the dividend is a corrective deduction for this structural injustice.

Crucially, the GRD is meant to be *structurally light*. It is not meant to require global government, nor a central taxing authority. It resembles a treaty-based regime, like the WTO or Montreal Protocol, agreed upon by states — with transparent enforcement through customs and reporting systems that already exist. And because the rate can be gradually adjusted, its distributive effects could scale over time as international coordination improves.

The GRD therefore crystallises Pogge's general strategy: small institutional adjustments in global property and exchange rules can produce massive improvements in human well-being without imposing severe burdens on the affluent. The GRD is meant to demonstrate that the eradication of extreme poverty is not blocked by material scarcity — only by the way we currently choose to organise the global economy.

The Health Impact Fund

A second proposal that Pogge has developed — again in collaboration with others — is the *Health Impact Fund* (HIF). The HIF focuses on a different dimension of global injustice: access to medicines. Pogge's target here is what he sees as a structural design flaw in the global pharmaceutical innovation system — namely the heavy reliance on patent-based pricing.

Under the current system, pharmaceutical companies are incentivised to set extremely high prices during the period of patent exclusivity in order to recoup research and development (R&D) costs and generate profit. This model can work for lifestyle drugs in affluent markets; but it systematically marginalises the poor. Patients in low-income countries, and in many middle-income countries, are often priced out of access to the very medicines that could prevent illness, disability or premature death. The result is again an institutional arrangement in which millions of people suffer avoidable harm — not because treatments do not exist, but because those treatments are commercially inaccessible.

Pogge's HIF seeks to flip this logic. Instead of rewarding companies for high prices, it would reward companies for *health impact*. The idea is relatively simple: pharmaceutical firms could voluntarily register a new medicine with the HIF. In return, they commit to selling this medicine worldwide at or near the cost of production. They forgo the "monopoly pricing" component of patent law — but receive compensation from a global public fund. That compensation is then allocated proportionally to the measurable health gains that the medicine delivers (for example DALY reduction — Disability-Adjusted Life Years). If a medicine saves many lives or prevents severe illness at scale, the reward to the company is correspondingly greater.

To work as a genuine alternative incentive structure, the HIF would require significant and stable funding — Pogge estimates around 6 billion USD annually — to be provided by participating states. But crucially, this would still be small relative to existing health budgets of wealthy countries, yet potentially transformative in its effects. The HIF would create financial incentives to innovate in areas currently neglected by profit-maximising firms — such as malaria, tuberculosis, and a large range of tropical diseases which impose heavy burdens on the global poor but promise very weak commercial returns.

There are, of course, challenges. Measuring health impact rigorously across contexts is not trivial. The HIF depends on sustained multilateral political commitment. And participation for pharmaceutical companies would remain voluntary — so firms might simply avoid it if the patent system continues to deliver higher payoffs.

Yet within Pogge's architecture of proposals, the HIF is important because it illustrates his normative shift: instead of merely telling individuals to give more (Singer), and instead of focusing only on moral condemnation, Pogge tries to design *institutional corrections that reshape the incentive landscape itself*. The HIF is meant as a proof-of-concept — a demonstration that we can redesign specific components of the global economic order in ways that greatly reduce avoidable suffering while imposing only modest costs on the affluent.

8.8 Rethinking Pogge's Critics

Although many philosophers share Thomas Pogge's moral urgency about global poverty, they diverge sharply on how to understand its causes, who bears responsibility for it, and what kind of moral duties follow from this understanding. Pogge's project represents one of the most ambitious attempts to extend the Rawlsian concern for justice beyond national borders. Yet, as several critics point out, his theory still inherits some of Rawls's limitations — an overemphasis on institutions, a narrow view of moral responsibility, and an abstract treatment of historical injustice. The critiques of Joshua

Cohen, Kok-Chor Tan, Neera Chandhoke, Leif Wenar, and Charles Mills, though distinct, converge around a set of common themes that reveal the boundaries of Pogge's approach (Jaggar, 2010).

A first major theme concerns the *causal role of global institutions*. Pogge's central claim is that the international economic order is not merely failing to reduce poverty but is actively producing and sustaining it. However, critics such as Joshua Cohen question whether global rules are truly the dominant factor. Some countries — notably China, Vietnam, and Bangladesh — have achieved rapid poverty reduction under the same global institutional conditions that have left others behind. This variation suggests that domestic policies, governance, and social institutions also play a crucial role. For Cohen, then, Pogge's "strong thesis" — that most poverty could be eliminated by minor global reforms — lacks empirical support. The causes of global deprivation are more complex, involving the interplay between international structures and local agency.

A second recurring criticism targets Pogge's reliance on *negative duties*. Pogge insists that our primary moral responsibility is to avoid participating in institutions that cause harm. In his view, citizens of affluent states are complicit in global poverty because they support or tolerate an international order that predictably inflicts deprivation on others. Critics such as Kok-Chor Tan and Neera Chandhoke argue that this minimalist framework is morally insufficient. Human rights, they claim, generate not only negative duties — to avoid causing harm — but also *positive duties* to assist those in need. Moreover, many humanitarian actors and NGOs are not motivated by guilt or complicity but by solidarity and compassion. Pogge's moral theory, centered on culpability, struggles to account for these universal moral impulses. Justice, they argue, cannot rest solely on the logic of harm; it must also rest on a broader conception of human moral concern.

A third point of contention concerns *the distribution of responsibility*. Pogge locates the main responsibility for reform in the citizens and governments of affluent countries, since they design and sustain the global order. Yet critics insist that this emphasis risks ignoring the duties of governments and civil societies in poorer nations. Chandhoke in particular stresses that domestic actors also bear first-order responsibilities to challenge injustice and build more accountable institutions. Global poverty is not simply imposed from above; it is also sustained by local elites who exploit international opportunities for corruption and rent-seeking. Assigning exclusive blame to the rich, she warns, can obscure the shared and interactive nature of moral responsibility in a globalised world.

Finally, several critics argue that Pogge's account remains *too abstract and too Rawlsian* in form. Leif Wenar, for instance, agrees with Pogge's diagnosis of the "resource curse" but rejects his proposed remedy — constitutional amendments and international democracy panels — as overly idealistic. Instead, Wenar proposes a more feasible framework grounded in existing legal and civil indicators, treating citizens as the rightful owners of their country's natural resources. Charles Mills takes the critique further. He argues that Pogge, like Rawls, still underestimates the historical depth of global injustice. The world order is not merely an unjust distribution of wealth; it is a *racialised and gendered* system shaped by centuries of colonial domination and patriarchal exclusion. For Mills, global justice must therefore begin as a non-ideal, *rectificatory* project — one that acknowledges race, empire, and gender as constitutive of the global basic structure itself.

Taken together, these critiques expose the philosophical and political limits of Pogge's global institutionalism. They suggest that a full account of global poverty must integrate both structural and local explanations, combine negative and positive duties, distribute responsibility across multiple levels of agency, and confront the historical legacies of racial and gender oppression. Pogge's framework remains one of the most influential attempts to theorise global injustice, but it is also a reminder that the moral architecture of global justice cannot rest on institutional design alone. To understand and address world poverty, we must move beyond Rawls's abstract universality and confront the concrete realities of power, history, and human interdependence that define the world as it is.

8.9 Conclusion: Plural Pathways Against Global Poverty

This chapter began from a stark empirical observation: despite decades of growth, innovation, and globalisation, hundreds of millions of people still live in conditions that systematically truncate their life chances. Extreme poverty is no longer the majority condition of humanity, but it remains a massive and avoidable deprivation, disproportionately borne by those least responsible for the rules under which the global economy operates. Against this background, the debate is not over *whether* global poverty is morally troubling, but over *how* it should be understood and *what* kind of response justice demands.

Our review of financial flows already complicates the conventional narrative of "rich countries helping poor countries through aid." Official Development Assistance (ODA) is large and politically prominent, but it is only one component in a wider ecology of cross-border resources. Remittances are larger in volume and more tightly linked to poor households; philanthropic funding is smaller, heterogeneous, and only partially poverty-focused. Empirically, aid clearly "does something": health-focused programmes have saved millions of lives, and humanitarian interventions can mitigate acute crises. At the same time, the broader macro- and institutional evidence remains ambivalent. Aid has neither produced the developmental miracles once promised, nor been the unambiguous disaster claimed by its harshest critics. Its effects are conditional: powerful in some domains, weak or counterproductive in others, and deeply mediated by domestic institutions.

This ambivalence underlies the competing assessments of Moyo, Deaton, and Sachs. For Moyo and Deaton, large-scale aid, as typically practised, tends to erode domestic accountability, entrench extractive elites, and crowd out the political bargains through which legitimate states are normally built. For Sachs, by contrast, the central problem is under-investment: extreme poverty persists because societies trapped at very low income levels cannot generate the surplus needed to invest; what is missing is a sufficiently ambitious and well-funded "big push." These diagnoses pivot on different understandings of what poverty fundamentally is: a shortage of capital (Sachs), or a failure of institutions and incentives (Moyo, Deaton). The resulting policy recommendations diverge sharply, but they share an important negative conclusion: simply scaling up “business-as-usual” aid will not be enough. Either aid must be transformed in its modalities (Sachs) or reined in and replaced by different forms of integration and finance (Moyo, Deaton).

In the normative register, Singer and Pogge reframe this empirical debate as a question of moral responsibility. Singer’s utilitarianism treats global poverty primarily as a vast aggregate of preventable suffering. On this view, affluent individuals are morally required to give, and to give effectively, until further sacrifice would impose comparable costs on themselves or their dependents. The moral emphasis falls on *what we can do now*: reallocating resources to highly cost-effective interventions, guided by evidence and impartial concern. Pogge, by contrast, locates the heart of injustice not in our failure to assist, but in our participation in institutional arrangements that foreseeably and avoidably harm the poor. Our most stringent duties, for him, are negative: not to uphold global rules that generate massive human-rights deficits. In this perspective, aid and philanthropy can at best play a compensatory role; the primary task is to reform the global basic structure – trade, finance, intellectual property, resource regimes – so that it ceases to produce deprivation in the first place.

Effective Altruism, as a practical outgrowth of Singer’s position, and Pogge’s proposals for a Global Resource Dividend, a Health Impact Fund, and the removal of the resource and borrowing privileges illustrate two very different strategies of moral institutionalisation. The former seeks to optimise the use of existing channels – philanthropy, individual giving, targeted programmes – under the current basic structure; the latter aims to modify that structure itself. Both approaches aim to reduce avoidable suffering, and both have attracted substantial criticism. EA is faulted for its reliance on a strong maximising utilitarianism and its relative silence on structural injustice; Pogge is criticised for overemphasising global institutions at the expense of domestic agency, for his narrow focus on negative duties, and for a somewhat abstract treatment of historical and racialised forms of domination. These critiques do not neutralise their contributions, but they do reveal that neither a pure “charity-plus-optimisation” model nor a pure “institutional harm” model can, by itself, capture the full moral and political complexity of world poverty.

What emerges from this chapter, then, is not a single master-solution, but a layered picture of partially overlapping, partially conflicting approaches. At the **empirical** level, the evidence suggests that aid is most defensible where it directly saves lives and supports clearly defined public goods, and most problematic where it substitutes for – rather than reinforces – domestic institutions and fiscal capacity. At the **normative** level, the debate oscillates between three foci: (1) *outcomes* (Singer, EA, Sachs), (2) *rights and harm* (Pogge), and (3) *agency and accountability* (Moyo, Deaton, many of Pogge’s critics). A plausible ethics of global poverty must find a way to integrate all three: to care about how many lives are saved, about whether the poor are wronged, and about who has the power – and the responsibility – to change the rules under which poverty persists.

In this sense, the chapter points toward a more plural and demanding conception of global justice. Reducing world poverty requires at least three kinds of action: (i) improving the effectiveness and accountability of existing aid and philanthropic instruments; (ii) reforming global economic rules that systematically disadvantage poor people and poor countries; and (iii) supporting domestic struggles for more inclusive and accountable institutions in the global South. These are not mutually exclusive agendas; they are interdependent. To focus only on efficient giving without addressing structural injustice is to risk managing symptoms indefinitely. To focus only on global rules while neglecting immediate, life-saving interventions is to tolerate preventable deaths in the name of long-term purity. And to shift all responsibility onto domestic actors in poor countries is to ignore the ways in which international structures shape their options.

The moral landscape of world poverty is thus neither simple nor hopeless. The analyses reviewed in this chapter show that there is no shortage of feasible reforms, or institutional imagination. What is lacking is not knowledge of *whether* we can do better, but political will over *which* burdens the affluent are prepared to accept, and *which* transformations of the status quo they are willing to pursue. Recognising global poverty as a matter of justice rather than misfortune implies that passivity is itself a form of complicity. The task for the remainder of this work is therefore not to ask whether we ought to respond to world poverty, but to clarify the terms on which that response can be ethically justified, politically realistic, and genuinely emancipatory.

Chapter 9

INEQUALITY AND MIGRATION

9.1 Introduction

The relationship between inequality and migration lies at the heart of the moral, political, and economic contradictions of globalization. The contemporary world is bound together by trade, technology, and communication, yet divided by borders that determine, with astonishing precision, who may access opportunity and who must remain excluded from it. Migration, in this sense, is not simply a demographic phenomenon — it is the human face of global inequality. It reveals how deeply the geography of birth continues to shape the trajectory of a person's life.

In an age of proclaimed universalism — where freedom, equality, and human rights are celebrated as global norms — the single most decisive factor determining one's life chances remains the country of origin. Economists such as Branko Milanović have shown that the majority of global inequality today is "between countries," not "within them." This means that the average income of the country into which one is born explains more about one's future prospects than one's effort, education, or class background. Citizenship, in this light, functions less as a moral identity than as an inherited asset — what Milanović calls a *rent-conferring privilege*.

Migration directly exposes this hidden structure of advantage. To move from a poor to a rich country is to cross not just a border but a moral threshold: it is to convert a low-value citizenship into a high-value one, to turn location itself into a form of capital. For millions, migration represents the most effective — and sometimes the only — path out of poverty. Yet for wealthy societies, it also raises fears about redistribution, cohesion, and identity. Immigration debates, therefore, are never simply about numbers or borders; they are about how societies define the scope of justice and the boundaries of solidarity.

This chapter begins from this tension. It examines how inequality is organized by geography and how states, through the control of movement, actively reproduce that geography. It explores why, in a world of open markets and global communication, human mobility remains so morally and politically contested. The question is not only empirical — how many people move, and with what effects — but also normative: what justifies keeping people apart when their movement could vastly reduce inequality and suffering?

The chapter unfolds by contrasting two great moral traditions of thought. On one side stand *cosmopolitan* theorists, who argue that borders have no intrinsic moral weight, that all human beings are equal in moral standing, and that restricting movement is tantamount to maintaining feudal privilege on a global scale. On the other side stand *communitarian* and *liberal nationalist* thinkers, who maintain that justice and solidarity are only possible within bounded communities capable of self-determination. For them, controlling membership is not a violation of justice but its precondition.

Between these poles lies a vast terrain of political tension. Cosmopolitan ideals confront democratic realities: even in societies that affirm human equality, electorates resist open borders, fearing economic strain, cultural loss, or erosion of social trust. Political feasibility thus becomes the testing ground for moral theory. Can a world that is morally committed to equality but politically structured by exclusion find institutional forms that mediate between the two?

The sections that follow trace this problem from its empirical foundations to its philosophical and policy implications. They begin by outlining the geography of global inequality — the scale of the “citizenship premium” — before turning to the major philosophical arguments that have shaped the debate, from Walzer’s communitarian defense of closure to Carens’s cosmopolitan case for open borders. The analysis then explores the search for “second-best” solutions — pragmatic institutional designs that might expand mobility without collapsing political support in host societies — and concludes by reframing the moral arithmetic of borders through Dani Rodrik’s thought experiment on how much more we must value our own citizens to justify keeping others out.

At stake is not only the future of migration policy, but the moral architecture of global justice itself. The border, in this account, is more than a line on a map: it is the visible expression of an invisible hierarchy, the point where the ideals of equality and the realities of exclusion meet. To understand inequality and migration together is to ask, in the most direct way possible, what it means to live in a world that recognizes every person as equal — but allows their place of birth to determine everything that follows.

9.2 Countries Inequality and Immigration

Imagine a world in which people could move across borders as easily as they move across administrative regions inside their own country. One could board a train or plane, show a passport, and enter another jurisdiction without visas, quotas, “waiting zones,” detention centers, or procedural suspicion. One could be expelled only if one is a genuine threat to security — not because one lacks the “correct” document. Is this a utopia? No: it already exists, but only for the citizens of wealthy countries. Western Europeans, North Americans, Japanese, Australians — and a handful of other rich polities — already live in this world. For them, borders are porous, airports are gateways, and temporary residence abroad is administratively easy. By contrast, for those who happen to be born in poor countries, borders are walls. If someone in Lagos, Dhaka, Kinshasa or Sana’a cannot find work to feed their family, they cannot simply leave. Their right to move — to maximise their own life chances through geography — is denied.

The empirical picture sharpens this reality. According to the most recent UN DESA revision (2023), there are roughly 281 million international migrants — about 3.6 percent of the world's population. In 2000, the figure was around 173 million, meaning that the migrant stock has risen by more than 100 million in twenty-three years, with annual growth of around 2–2.5 percent — almost twice the rate of world population growth. This is the realised flow: the people who actually managed to cross and stay. But aspiration surveys such as Gallup's global polling suggest the latent demand is much larger: around 900 million adults — roughly 15 to 17 percent of the world's adult population — say they would like to move permanently to another country if they could. The sizes of immigrant populations in countries like the United States (about 14 percent foreign-born) and Spain (about 15.5 percent) show that these are not utopian numbers — when permitted, these shares materialise.

Why is this so? Because the global system attaches enormous value to *location* — to citizenship.

Recent empirical work on global inequality (Milanović (2016)) shows that most of the inequality in the world is still explained by differences *between* countries, not by differences *within* them. The "location" component — the part of income inequality determined by the average income level of the country where a person lives — increased strongly during the 19th and early 20th century, peaked around 1975, and although it has fallen slightly in the last decades (mainly because of rapid growth in China), it still accounts for the largest part of global inequality today.

We can think of global inequality as the sum of two parts:

global inequality = inequality between countries (location) + inequality within countries (social class)

We can clarify the difference between the "location" component and the "social class" component with a simple thought experiment.

Imagine an unborn person who is about to enter the world. We do not yet know who they will be. Two pieces of information will determine his position in the global income distribution:

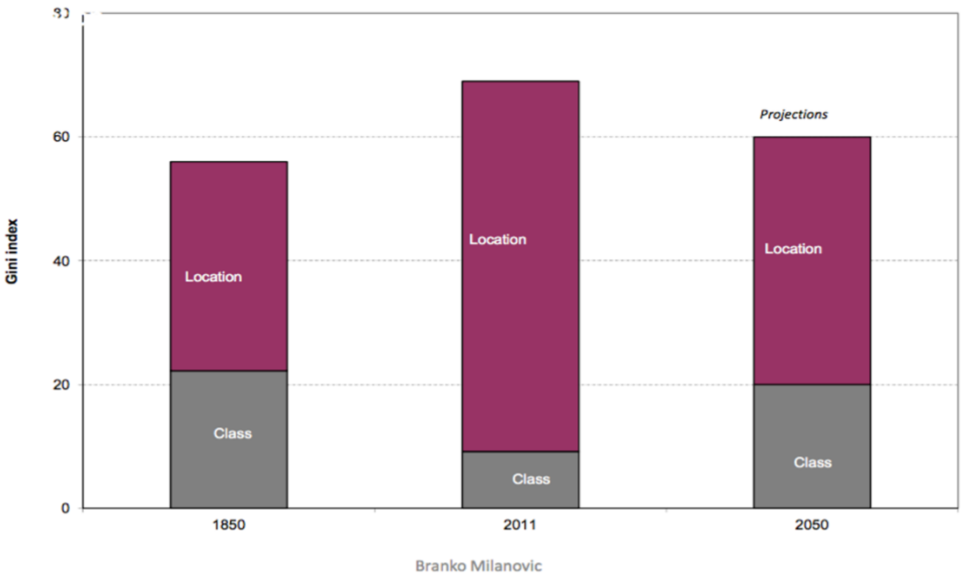
1. **the country in which they will be born**
2. **the social class they will belong to within that country**

Now ask: which of these two factors matters more for predicting his lifetime income?

The answer is: *country*.

Because income differences between countries are so large, the "location" factor dominates. A poor person in a rich country (for example, in the bottom part of the US income distribution) is usually much better off than a middle-class or even upper-middle-class person in a very poor country. The gap between national averages is so vast that “class” differences inside countries come second.

In other words: the biggest determinant of a person’s place in the global income hierarchy is not whether they will be born into a poor or rich family — it is whether they will be born into a poor or a rich *country*. This is exactly what Milanović means when he says that global inequality is largely driven by the “location” component.



This becomes even clearer when we look at mobility: **97% of the world’s population still lives in the country where they were born**. In practice, people do not move. As a consequence, most differences in people’s life chances reflect the income level of the country they happen to be born into, not differences in personal effort or social class within that country.

Country	Citizen premium index (US = 100)	
United States	**100,00**	X
Germany	86,80	1,2
Sweden	85,50	1,2
France	77,60	1,3
United Kingdom	76,30	1,3
Italy	65,80	1,5
Spain	63,20	1,6
Brazil	23,70	4,2
Mexico	26,30	3,8
Argentina	30,30	3,3
Egipt	20,40	4,9
Morocco	18,40	5,4
Algeria	19,70	5,1
Nigeria	7,90	12,7
Kenya	6,60	15,2
Ethiopia	3,90	25,6
DR Congo	1,40	71,4
China	32,90	3,0
Philippines	31,60	3,2
India	11,80	8,5
Indonesia	19,10	5,2

Values based on approx. 2023 GNI per capita PPP.

The income gain associated with being born in a rich country therefore functions like a rent. If we take the Democratic Republic of Congo as a baseline, the average income in the United States is around 71 times higher, in Sweden around 62 times higher, and in Brazil around 17 times higher. In other words, someone born in the United States has more than 70 times the income potential of someone born in Congo — purely because of their location at birth.

Milanović therefore calls citizenship a “rent-conferring privilege”: a pure return to birthright, not to effort or merit. Citizenship is binary. You have it or you do not. And this binary determines the scale of your life chances.

Precisely because the passport premium is so large, wealthy states invest heavily in preventing entry. Walls, fences, minefields, maritime interceptions, fortified borders — these are not metaphors but physical infrastructures. The United States maintains steel fencing and patrols across almost two thousand miles of the US–Mexico frontier. The European Union finances Libyan coast guard interception to prevent Africans from reaching Lampedusa or Sicily. Israel operates a concrete barrier. Saudi Arabia does so along its border with Yemen. The Korean peninsula is divided by a demilitarised zone planted with mines. India fences Bangladesh. EU external frontiers in Bulgaria or Poland have begun erecting anti-migration fences. In short, the contemporary global order permits radical mobility for goods, finance and technology — but not for the human beings who produce those goods, finance, and technologies.

The world we inhabit is therefore not one of equal opportunity but of *territorially protected rents*. What appears as “border control” is, in distributional terms, the defence of citizenship privilege. The right to move freely already exists — simply not for everyone.

Cosmopolitans such as Pogge, Beitz, and Singer argue that in an interdependent world we cannot treat large differences in life chances between nations as morally inert. When the **citizenship premium** is as large as it is, inequalities of birthplace become hard to defend from a liberal point of view. The poorest Americans are, by global standards, relatively well-off. Within domestic societies we view inherited advantage with suspicion: we tax intergenerational transfers, we implement policies to equalise schooling and access to opportunity. Yet when confronted with two individuals born in different countries, we do not even begin to consider whether they should have **equal opportunities** across borders. Rawls’s decision to terminate the scope of equality of opportunity at the national boundary is precisely what cosmopolitans reject. For them, if the world is economically interdependent, then the relevant moral relationships are not exhausted by state-to-state diplomacy – individuals are already connected directly through global production chains, market relations, and institutional externalities. A form of implicit social contract must therefore be understood to exist among citizens of the world.

There are only three structural channels through which global inequality can be reduced: (1) the acceleration of income growth in poor countries; (2) global redistributive mechanisms; and (3) migration. Official Development Assistance today is a little above 200 billion dollars per year – roughly 0.10% of world GDP. The magnitude of this transfer is trivial relative to the scale of global inequality. Thus, the choice becomes stark: *either* poor people become richer where they are, via domestic growth or international redistribution, *or* they become richer by moving to places where income is already high. Migration is therefore not a marginal curiosity: in a world where location determines most of the variance in life outcomes, it is a direct instrument for appropriating the **citizenship premium**. Although only 3% of the world population are migrants, remittances - as it has been highlighted in the previous chapter - are the largest external resource for poor households. In the Philippines remittances account for roughly 9-10% of GDP, one of the highest ratios in the world for a major economy.

There is, then, a fundamental contradiction at the heart of contemporary globalization. We treat capital, goods, and technology as factors that should flow freely across borders in pursuit of efficiency and welfare gains — but we restrict the mobility of labour, even though labour mobility is precisely the channel through which individuals can most directly appropriate the **citizenship premium**. As it has been pointed out by the economist, Branko Milanovic, this contradiction generates four specific tensions.

First, people are increasingly willing to leave — and yet they lack any corresponding right to move where they see fit. Second, globalization doctrine demands the free movement of factors of production, but the one factor that would do most to equalize global returns — labour — remains politically constrained. Third, the maximization of income at the global level presupposes that each economic agent should not be bounded by the arbitrary limits of nation-state jurisdiction. Fourth, development theory emphasizes improving people's lives within their home countries, but a broader conception of development would consider improving lives wherever they can be improved — regardless of where a person currently resides.

This is where the inconsistency with free trade becomes especially visible. Free trade is justified on the grounds that, although some domestic groups will lose, the overall social product is higher; and once this higher total income is achieved, redistribution can be undertaken — in principle — to compensate those who are worse off. Yet we do not apply the same logic to migration. When the global social product could be maximized by allowing people to move, we refuse to allow the movement, and we do not even discuss ex ante transfers as a compensatory scheme. We accept the Kaldor–Hicks logic for goods and capital flows, but we suspend it when the relevant factor is human mobility. The question that then emerges is straightforward: *why is free movement of goods permissible, when free movement of the human beings who produce those goods is not?*

The usual explanation is not theoretical consistency, but political asymmetry. Free trade does not alter the composition of the political community — whereas migration does. Goods do not vote, do not enter schools, and do not claim social rights; people do. For this reason states fear that open borders would transform the demos, create fiscal obligations, and disrupt domestic political coalitions. At the same time, globalisation has built strong international institutions for the mobility of capital and goods, but not for the mobility of persons. The result is a structurally asymmetric order: markets are global, membership remains national. Cosmopolitans argue that this asymmetry between free movement of goods and the restricted movement of people is unjustifiable because it protects the interests of the already advantaged. Free trade allows wealthy countries and firms to extract value from cheap labour abroad, while preventing those same workers from accessing the higher returns available inside rich economies. In other words, rich states benefit from global markets when it suits them (as consumers and investors), but they close borders when it would benefit the poor (as potential migrants). The principle of equal moral worth requires that persons — not commodities — should be primary. Allowing goods to circulate freely while preventing the human beings who produce them from moving simply entrenches location-based privilege and preserves the citizenship premium — and cosmopolitans argue that there is no coherent moral justification for that.

9.3 The Philosophical Case for Exclusion: Walzer, Wellman, Miller

In contrast to cosmopolitan defenders of open borders, a significant line of argument in contemporary political philosophy insists that states have a moral right to restrict immigration. This tradition begins with communitarian thinkers, especially Michael Walzer, and continues through a later liberal-nationalist strand represented by Christopher Heath Wellman and David Miller. They all reject the idea that borders are morally arbitrary. But they do so on different grounds. Their arguments illuminate not only *why* borders matter, but also what kind of community a political society is.

Michael Walzer's view, developed in *Spheres of Justice* (1983), is the classic communitarian defence of closure. For Walzer, justice is not a universal procedure applied "from nowhere," but something internal to historically formed communities with shared meanings, shared understandings, and shared practices. Before one can distribute social goods fairly, one must first determine who the recipients of those goods are. Membership precedes distribution. As he famously puts it, democratic equality is only possible behind at least partially closed borders, because without the power to decide who belongs, no common life — and hence no egalitarian distribution — can be sustained. Immigration policy is therefore not primarily about labour markets or efficiency, but about preserving the moral identity and integrity of a political community. On this account, "open borders" is in a deep sense incoherent: a community that cannot determine its membership cannot meaningfully determine its own values.

Christopher Heath Wellman's defence of immigration restriction is grounded not in cultural particularism, but in a liberal principle that most people would otherwise regard as uncontroversial: *freedom of association*. In his influential article "Immigration and Freedom of Association" (Wellman 2008), he argues that if we take seriously the right of persons to decide whom to associate with — and equally, whom *not* to associate with — then we must, to be consistent, grant political communities a similar moral right. For Wellman, states are not merely administrative corporations; they are legitimate collective agents constituted by their members. And like individual moral agents, they possess rights that attach to the exercise of their self-determination. Among these rights is the freedom not to admit outsiders.

This argument is deliberately analogical. Individuals choose whom to marry, whom to live with, whom to invite into their home, and whom to join in voluntary clubs. In all of these domains, we acknowledge a right of refusal: one may decline friendship, decline proposals, decline to live with someone, decline membership. No one is morally compelled to accept every possible partner merely because that partner desires entry. Wellman extends this logic: if it would be absurd to say that an individual is morally obliged to welcome all strangers into his or her home, it is likewise implausible to say that a political community is obliged to welcome all would-be migrants into its territory.

This is why Wellman resists attempts to label border controls as inherently xenophobic. The right to exclude does not need to rely on racial hierarchy, cultural chauvinism, or civilizational prejudice. The right arises instead from the structure of liberal rights themselves. A state may permissibly wish to preserve its capacity for collective self-determination — its capacity to legislate, tax, educate, regulate, and pursue its political projects — without thereby morally maligning those outside. Exclusion, on this view, is not a judgement on the value of the outsider; it is a protection of the autonomy of the group deciding its own future.

It is important to see, however, that Wellman's defence is not absolute. He acknowledges that the right to exclude cannot trump *all* other moral duties. In cases of urgent humanitarian peril — refugees who face death, genocide, or catastrophic threat — a state's presumptive freedom of association must yield to the more fundamental obligation not to stand by while innocent persons are destroyed. Likewise, basic human rights constraints continue to apply: a state cannot violate refugees' bodily integrity, or torture or kill them, simply because it wants to protect its associative autonomy. So while Wellman defends a general right to exclude, this right exists within a moral framework that makes room for enforced admission in cases of humanitarian emergency.

The result is a distinctive liberal nationalist position: borders are morally meaningful, not because shared culture must be defended (as in Walzer), nor because national identity is intrinsically valuable (as in some communitarianism), but because free association implies a corresponding right to non-association. The moral core, for Wellman, is not ethnic particularity but liberal self-determination.

David Miller's position represents perhaps the most systematic liberal-nationalist defence of immigration restrictions. In *Immigration: The Case for Limits* and later in *Strangers in Our Midst*, Miller argues that states may permissibly restrict admission not out of xenophobia or cultural chauvinism, but because nations are legitimate collective agents with special distributive obligations to their own members. For Miller, national communities are not accidents of geography. They are historically developed frameworks of solidarity and trust — and because egalitarian redistribution depends on this solidarity, the stability of the national community has direct normative weight.

The central core of Miller's argument is national self-determination. If individuals are entitled to control the direction of their lives, democratic peoples — conceived as self-governing collectives — must likewise have a right to exercise control over their future. Borders are part of this. Who is admitted and who is not has long-term consequences for the cultural trajectory, social trust, and institutional cohesion of the political community. Unlike cosmopolitans — who see this as morally irrelevant — Miller insists that national culture is the background condition that allows egalitarian institutions such as the welfare state to function. If immigration is so large or unmanaged that it reshapes the cultural basis of trust faster than the public can absorb, then the motivational foundations of distributive justice risk collapsing. A national welfare state depends on citizens believing that "we" owe something to "us."

This is why Miller rejects both open borders and total exclusion. He argues instead for "moderate limits": enough control to preserve the continuity of national identity and the sustainability of social cooperation, but not so much restriction as to seal the nation off from humanitarian responsibility or economic dynamism. The emphasis is equilibrium — not closure.

Distributive justice plays a key role. Miller holds that states have primary duties to their own citizens, particularly those who are economically vulnerable. Since large inflows of low-skill migrants may depress wages at the bottom of the labour market or strain welfare systems, governments have a morally legitimate reason to limit immigration in order to protect their poorest citizens. Equal regard, on Miller's account, is not cosmopolitically global — it is structured through national membership.

Yet Miller insists that these rights are not absolute. States cannot ignore the plight of refugees facing persecution or mortal danger. Human beings in extreme peril may have a claim that overrides the usual discretionary right to exclude. But even then, Miller argues that rescue does not necessarily entail full membership or immediate citizenship. Humanitarian assistance can be fulfilled through temporary protection, regional solutions, or burden-sharing agreements.

Miller's position therefore occupies a nuanced middle ground. Against cosmopolitan theorists, he denies that borders are morally arbitrary. But unlike hard communitarians, he does not claim that cultural preservation can justify any means. His defence of controlled borders rests on the conditional value of national community as the enabling condition of egalitarian justice. Citizenship, for Miller, imposes special duties — and immigration policy is one expression of how those duties can legitimately be prioritized.

So while Walzer, Wellman, and Miller converge in rejecting open borders, they do so for strikingly different reasons. For Walzer, community is a moral world with its own internal logic; admitting newcomers is not only a demographic act, but a transformation of the very meaning of the common life. For Wellman, the justification is thinner but more general: the right to exclude is simply the collective analogue of an individual's right not to associate. For Miller, the justification is "national" rather than communitarian or liberal: a nation must protect the social trust that sustains egalitarian redistribution. All three disagree sharply with cosmopolitans like Carens and Pogge on whether citizenship can be treated as a morally arbitrary border. And yet even here they differ among themselves about *what* is being protected: a cultural form of life (Walzer), a right (Wellman), or a redistributive infrastructure (Miller).

Far from being a single unified doctrine, the philosophical defence of borders is a family of arguments with different normative centres. The communitarian sees a community of shared meanings; the associational liberal sees the structure of rights; the liberal nationalist sees a collective good that must be protected. The same policy (limiting migration) is thus justified by three different moral logics.

9.4 Cosmopolitan Arguments for Open Borders: Singer, Pogge and Carens

Cosmopolitan theorists reject the idea that the moral community stops at the state border. If persons are what ultimately matter, then whether those persons happen to be born inside or outside our nation cannot, on its own, have fundamental moral significance. Two of the most influential contemporary philosophers who articulate this approach — although neither has written a single monograph on migration the way Joseph Carens has — are Peter Singer and Thomas Pogge. Both reach the defence of expanded mobility not through communitarian premises, but through their more basic commitments: for Singer, utilitarianism; for Pogge, global institutional moral duties.

Peter Singer never wrote a full systematic treatise on immigration, but the logic of his work — from *Famine, Affluence and Morality* (1972) to *One World* (2002) and *The Life You Can Save* (2009/2019) — leads directly to the conclusion that preventing people from moving to where their lives would go markedly better is morally indefensible. Singer's utilitarianism does not recognise borders as morally relevant dividing lines. The good to be maximised is *global well-being*; suffering is suffering wherever it occurs. It is therefore inconsistent — in Singer's view — that we celebrate the free movement of goods and capital in the name of efficiency, yet refuse free movement of the human beings whose welfare is directly at stake. In *One World* (2002), he makes this point explicitly: if we really care about improving global welfare, then labour mobility should be seen as a more direct and more powerful instrument than trade liberalisation, because moving the person itself, not the product, generates welfare gains on the order of 2x, 5x, even 10-fold. When the gains to the migrant are this dramatic, Singer argues, immigration restrictions amount to the wealthy blocking enormous benefits to the poor for the sake of protecting relatively minor advantages to themselves. On his framework, the only consistent position is to relax border controls significantly.

Thomas Pogge arrives at a similar conclusion through a different route. Pogge's central theme — especially in *World Poverty and Human Rights* (2002; revised 2008) — is that severe global poverty is not merely an unfortunate natural fact, but the foreseeable result of a global institutional order crafted by affluent states. If the global rich are partially responsible for sustaining an economic system that predictably reproduces massive inequality in life prospects between countries, then they owe compensatory duties to those whom the system harms. Much of Pogge's writing focuses on trade rules, resource privileges and the WTO structure, but migration is woven into this as both a symptom and a potential remedy. In several essays (e.g. Pogge 1997; 2005), he notes that affluent societies simultaneously sustain the structures that impoverish the global poor *and* then close their borders to the very people whose opportunities have been damaged by this system. Nothing illustrates the injustice of the current order more dramatically, in his view, than this combination: first contribute to making people poor, then deny them the right to escape this imposed disadvantage. For Pogge, then, liberalising migration is not simply an optional humanitarian gesture — it is partially a matter of rectification. Giving people access to the labour markets of the rich world is one direct way of undoing some of the harm that affluent societies help to create.

Singer and Pogge thus begin the cosmopolitan case from two different ends — global utilitarianism in Singer's case, global institutional justice in Pogge's — but converge upon the same normative conclusion: the very large welfare gains that migration would generate for the global poor, combined with the relative moral weakness of the citizen-based claims used to block those gains, make immigration restrictions extremely difficult to justify from a global moral point of view.

Joseph Carens is the contemporary philosopher who most explicitly develops the claim that immigration restrictions are not merely unfortunate or suboptimal, but *deeply unjust*. Across several seminal essays — especially "Aliens and Citizens: The Case for Open Borders" (1987) (see Box # 1 for an exposition of the article) and *The Ethics of Immigration* (2013) — Carens argues that the freedom of movement across borders is not a luxury or a charitable concession, but something that follows directly from the moral foundations of liberal democracy itself. His central move is to reverse the burden of justification: in a liberal egalitarian normative framework, the *default* position should be open borders; therefore it is not proponents of open borders who must prove why people should be free to move, but restrictionists who must prove why it is legitimate to stop them.

Carens begins from the principle of the *moral equality of persons*. Liberal democracies already endorse — at least internally — the claim that individuals are moral equals regardless of race, class, religion, language, or gender. Freedom of movement is already treated domestically as a basic liberal right. A citizen may freely move from Milan to Naples, from California to New York, or from Lyon to Marseille. If one asked a liberal democrat why freedom of movement within the national territory must be respected, she would answer: because individuals should not have their life chances arbitrarily constrained by local authorities. Carens' question is simply: why does this argument magically stop at the border? Why is it morally unjust for Piedmont to prohibit someone from moving from Calabria, but perfectly acceptable for a State prohibit a foreigner to enter the country? The border does not appear, from the point of view of the moral person, to have intrinsic normative force. A border—simply as a geographical or political line—does not by itself generate moral duties or moral distinctions.

This leads to Carens' most famous analogy: the comparison between modern border controls and *feudal privilege*. In feudal Europe, people's opportunities in life were determined almost entirely by the social class and the estate into which they were born: serf, artisan, noble. These were *status-based systems of exclusion*. Carens argues that nationality functions today as a structurally equivalent mechanism:

whether a person is born in Stockholm or in Sierra Leone determines more than almost any individual factor — education, talent, work ethic — how that person's economic life will unfold. This is exactly the Milanović "citizenship premium" argument before Milanović. Nationality is the new feudal birth-status. And since liberalism is committed to the abolition of birth-ascription systems — we no longer think it acceptable to assign advantages and disadvantages based on caste or lineage — it should logically reject the moral legitimacy of national closure. For Carens, restricting access to membership in rich states is the contemporary functional analogue of medieval lords keeping serfs tied to the land. Borders are morally arbitrary.

Carens is also explicit that immigration restrictions cannot be excused by the appeal to distributive justice within one's own national community. If liberal egalitarians believe that justice requires compensating for morally arbitrary inequalities — like being born into a rich or poor family — then the *largest* morally arbitrary inequality of all — being born into a rich or poor country — cannot simply be ignored because it is inconvenient. Carens' cosmopolitan egalitarianism therefore sharpens the Pogge–Singer line: if we take liberal egalitarian justice seriously, open borders are not a utopian fantasy but the logical extension of our own principles. The domestic logic of freedom of movement and equal moral status commits us to free movement globally — unless one can demonstrate extraordinarily powerful countervailing reasons.

Finally, Carens emphasises that migration is not merely an abstract moral puzzle, but a direct policy mechanism for reducing global inequality. Given the empirical reality that moving across borders can multiply an individual's income by factors of five, ten or more — the "citizenship rent" — preventing migration is effectively preventing the world's poorest from accessing the only intervention that reliably and predictably transforms their life prospects. To Carens, then, the contemporary border regime is morally indefensible precisely because it seals up the wealth of the rich behind the accident of birthplace. Modern border controls are a closed aristocracy at planetary scale.

Carens's conclusion follows with clarity: if liberal democratic states sincerely believe in moral equality of persons, freedom of movement, and the rejection of birth-ascribed privilege, then they should — by their own principles — presume in favour of open borders. Restrictions are not the baseline. They are the anomaly that requires justification.

Box # 9.1 - Carens, Liberal Moral Theory, and the Ethics of Immigration Control

This box draws on Joseph H. Carens's seminal essay *"Aliens and Citizens: The Case for Open Borders"* (1987), which has become a foundational text in contemporary debates on the ethics of immigration. Carens argues that the conventional view—according to which states possess a sovereign prerogative to control immigration—is difficult to reconcile with the core moral commitments of the liberal political tradition. By systematically examining the implications of leading liberal theories, including utilitarianism, libertarianism, and Rawlsian contractualism, Carens demonstrates that these frameworks converge on a set of conclusions broadly supportive of open borders. Only communitarian perspectives offer a principled defense of restrictions, and even these, Carens contends, rely on assumptions that are difficult to sustain within a liberal moral culture. The following analysis reformulates Carens's argument in a systematic way, highlighting the internal logic of each moral theory and its implications for the ethics of migration.

Borders are typically enforced through the deployment of surveillance, policing, and legal sanctions designed to prevent individuals from entering a territory without authorization. Although such measures are often justified on the grounds of national security or the need to deter criminal activity, the overwhelming majority of prospective migrants are peaceful individuals seeking employment, safety, or improved life prospects. This raises a foundational normative question: **on what moral grounds may such persons be excluded?**

While the sovereign state traditionally claims the right to admit or exclude non-members, Carens argues that the central moral theories within the liberal tradition—utilitarianism, libertarianism, and contractualism—each undermine this conventional assumption. Taken seriously, these theories converge on the conclusion that immigration restrictions are difficult to justify. Only communitarian approaches support a robust right to exclude, and even they are subject to significant internal tensions.

Libertarianism: Property Rights, Voluntary Exchange, and the Limits of State Authority

Libertarian theory, exemplified by Nozick, grounds political authority in the protection of natural rights, especially individual property rights and the freedom to engage in voluntary exchange. Under this framework, a prospective migrant does not violate any natural right by crossing an international border, provided they do not trespass on privately owned land. The state's legitimate function is confined to enforcing pre-existing rights, not to restricting consensual economic interactions.

If, for instance, an American employer wishes to hire a Mexican worker, state intervention constitutes a violation of both parties' liberty. Competitive disadvantages to domestic workers do not generate a claim to protection, since no individual has a right to be shielded from market competition. Individuals may refuse to transact with non-citizens, but they cannot prevent others from doing so.

Consequently, the nationalist claim *"It is our country; we may exclude whomever we wish"* is incompatible with libertarian principles. From this perspective, immigration restrictions amount to an unjustified limitation on individual freedom. **Libertarianism, therefore, supports open borders as a direct implication of its core commitments.**

Contractualism: Rawls, Cosmopolitan Extensions, and the Original Position

Rawlsian contractualism endorses an active state entrusted with safeguarding basic liberties and promoting social justice. Rawls himself offers only limited commentary on migration, suggesting in *The Law of Peoples* that states may regulate immigration to maintain public order or protect environmental and demographic stability. However, these narrow considerations do not undermine the broader logic of Rawlsian justice.

Cosmopolitan theorists extend the **original position** globally. Behind the veil of ignorance, individuals lack knowledge of their national citizenship—an attribute Rawls characterizes as morally arbitrary. Reasoning under such conditions, the parties would recognize freedom of movement as integral to the pursuit of one's life plans and thus as a basic liberty. They would also reject immigration restrictions based on birthplace, culture, or inherited political membership. In the non-ideal world, characterized by vast inequalities and authoritarian regimes, temporary restrictions might be justified. This is the case, if migration would lead to chaos and thus would be a threat to basic liberties. In this case according the contractarian view is legitimate to limit immigration.

Yet these exceptions confirm the principle: restrictions require stringent justification, not mere appeals to national preference or cultural preservation. Rawlsian contractualism, especially in its cosmopolitan elaboration, thus **supports a strong presumption in favor of open borders.**

Utilitarianism: Equal Consideration and Global Welfare Maximization

Utilitarianism rests on the principle that each person's welfare counts equally in the moral calculus. Because citizenship is irrelevant to moral worth, the interests of migrants—and of individuals in their countries of origin—must be included alongside those of citizens in receiving states.

Given the magnitude of global inequality, the welfare gains that migrants obtain by moving from poor to wealthy societies vastly exceed the potential losses experienced by citizens. Even when some domestic workers face increased competition, these costs are typically outweighed by:

- substantial benefits to migrants and their families,
- positive spillovers to sending states, and
- economic gains in host societies.

While utilitarian conclusions depend on empirical assessments, the enormous potential welfare gains for migrants make it highly unlikely that restrictive immigration policies would maximize overall utility. Utilitarianism therefore **tends toward an endorsement of much more open borders.**

Communitarianism: The Main Source of Justification for Exclusion

Communitarian theorists, such as Walzer, present the most influential argument for state control over immigration, grounding it in the right of political communities to preserve their distinctive cultures and political institutions. On this view, membership is a collective good that citizens may regulate.

Yet even Walzer acknowledges important constraints:

- states must provide aid to the severely needy;
- once admitted, non-citizens must be offered full membership;
- large, sparsely populated states have weaker claims to exclude.

Carens argues that these concessions reveal the extent to which communitarianism is shaped by liberal commitments. Moreover, examples such as free mobility among U.S. states or within the European Union demonstrate that cultural variation is not incompatible with open internal borders. If freedom of movement can override local claims within a federal or supranational union, the moral distinctiveness of national borders becomes increasingly difficult to defend.

Across the major strands of liberal moral theory, the convergent conclusion is clear:

- **Libertarianism** supports open borders to protect individual liberty and property rights.
- **Contractualism** supports open borders because citizenship is morally arbitrary and freedom of movement is essential for pursuing a life plan.
- **Utilitarianism** supports open borders because they generally maximize global welfare.

Communitarian arguments offer a principled basis for exclusion, but only by appealing to values that sit uneasily with the deeper liberal commitments to moral equality, freedom, and impartial justification. As Carens concludes, restrictive immigration policies in liberal democracies stand in significant tension with the theories and values that shape those societies. The ethical logic of liberalism points decisively toward a presumption of open borders.

9.5 The Moral and Economic Case for Opening Borders: Efficiency, Equality, and Justice

Allowing workers from poor countries to move to rich countries would make the world vastly more equal — far more than trade in goods or capital ever could. The gains from labour mobility are not only economic in magnitude but moral in nature. Both utilitarian and egalitarian frameworks, though built on distinct foundations, converge on this point. For **utilitarians like Singer**, the argument rests on welfare maximisation: migration yields extraordinary gains in well-being at minimal cost. For **egalitarian theorists like Pogge**, it rests on justice: migration rights serve as partial redress for an international order that entrenches poverty by design. Together, these perspectives reveal that opening borders is not a moral luxury, but one of the most direct and defensible instruments for reducing global inequality.

Utilitarian Efficiency: Maximising Global Welfare

From a utilitarian standpoint, the principle is simple: all suffering and well-being count equally, regardless of nationality or geography. If a Bangladeshi worker earning $500 a month can earn four times as much by performing the same work in a higher-productivity economy, preventing that move inflicts an avoidable moral loss. In *One World* (2002), Singer highlights this inconsistency: the same societies that celebrate the free flow of goods and capital — in the name of global efficiency — erect barriers against the free movement of people, whose welfare gains would be exponentially greater.

As Rodrik (2018) demonstrates, even when societies place several times more moral weight on the welfare of their own citizens than on foreigners, relaxing immigration restrictions still increases total welfare. Rodrik's argument will be presented more in details in section 9.7. The logic behind opening up to immigration mirrors that of trade: the "size of the pie" expands, even if distributional adjustments occur within. Yet because barriers to labour mobility are vastly higher than tariffs on goods, lowering them even slightly generates disproportionately large efficiency gains. In short, migration liberalisation offers far more "bang for the buck" than trade liberalisation**.**

Moreover, the fear that immigration undermines labour standards is largely misplaced. When rich countries import goods from low-wage economies, they often externalise exploitation: firms relocate to where wages and safety standards are lowest — what critics call *social dumping*. By contrast, when workers move to rich countries and work under domestic labour laws, they are covered by minimum wages, safety protections, and the right to organise. Properly regulated migration does not erode standards; it globalises them. Labour mobility, in this sense, is a corrective to the very inequalities that trade can reinforce. It raises efficiency without triggering a race to the bottom.

Thus, from a utilitarian welfare perspective, the conclusion is clear: migration generates immense aggregate gains, improves individual lives by orders of magnitude, and, unlike unregulated trade, can do so without dismantling domestic protections. It is the most direct and humane way to increase global well-being.

Egalitarian Justice: Rectifying Structural Inequality

From an egalitarian-rights perspective, the argument takes a different path but reaches the same destination. Thomas Pogge's analysis of global justice begins not with efficiency, but with complicity. The inequalities between nations, he argues, are not natural facts; they are the predictable result of a global institutional architecture shaped by and for the affluent. Trade regimes, resource privileges, and financial governance systematically advantage the citizens of rich countries. Within such a system, to restrict migration is not to preserve order, but to **preserve the spoils of an unjust order**.

On this view, labour mobility becomes an act of moral repair. If the international economic system contributes to making people poor, rich societies incur duties not of charity but of rectification. Preventing migration is morally equivalent to blocking escape from harm that one has helped to inflict. The right to move, therefore, is not a discretionary humanitarian favour extended by states, but a partial restoration of justice owed to those who suffer the structural consequences of an inequitable global order.

Where Singer sees migration as the most efficient route to maximise total welfare, Pogge sees it as a concrete instrument to dismantle the institutional roots of inequality. Both positions converge on the same empirical truth: **mobility is the most powerful equaliser in the contemporary world economy**. To deny that mobility is to preserve, by omission, a global caste system in which citizenship determines life chances more rigidly than any domestic class hierarchy.

Convergence and Political Feasibility

The utilitarian and egalitarian arguments are, in the end, complementary. The former stresses the scale of potential welfare gains; the latter the moral urgency of redress. Together they form a dual foundation for the same conclusion: that borders, as currently enforced, represent not just inefficiency but injustice.

The practical question, of course, is political. Voters in rich democracies tend to interpret migration not as global justice, but as domestic redistribution — a transfer of scarce privileges from "us" to "them."

The broader insight remains: whether viewed through the lens of welfare maximisation or of egalitarian rights, migration is both the most efficient and the most just mechanism for reducing global inequality. Opening borders, even partially, does not require utopian idealism. It simply requires consistency — treating human mobility with the same moral and economic logic we already apply to goods, capital, and ideas.

9.6 Cosmopolitanism vs. Communitarianism: a structurally unresolved conflict — and the search for "second-best" mediating designs

The disagreement between cosmopolitans and communitarians/liberal nationalists on immigration is not merely an academic quarrel over conceptual priority; it marks one of the deepest normative fractures in contemporary political philosophy.

On one side, cosmopolitans (Pogge and Carens) argue that borders are morally arbitrary; nationality is not a desert-relevant property; and the interests of a Congolese agricultural worker count just as much as the interests of a low-skilled American warehouse worker. Under this view, preventing people from moving to where their labour is worth more is not just economically inefficient — it is *unjust*. It is equivalent to blocking access to life-chances on the basis of birth. Moral equality, on their account, demands much more than charity; it demands the dismantling of structural privileges that are baked into the geography of citizenship. Singer as utilitarian would simply say that open borders will improve global welfare.

Communitarians (Walzer) and liberal nationalists (Wellman, Miller) fundamentally reject this starting point. Political obligation, they argue, *begins* inside a bounded political community — not across humanity as a whole. States, as self-determining associations, have special duties to their members *first*. They may wish to show humanitarian concern to outsiders, but they are not under a general obligation to share citizenship, welfare institutions, or labour markets with all human beings. This is not xenophobia on their own terms; it is simply the logic of bounded self-government.

These two starting points do not converge. They are structurally different pictures of political morality.

And crucially: **the political system reacts**.

In actual democratic politics, electorates in high-income democracies overwhelmingly exhibit communitarian preferences. Borders — for them — are not just administrative devices. They are identity markers, welfare boundaries, and psychological anchors for solidarity. When cosmopolitan arguments are pushed directly, the political response is not incremental liberalisation, but *backlash* — welfare chauvinism, border securitisation, externalisation agreements, and reactively tighter visa regimes. The last decade of EU, UK, US, and Australian practice confirms this pattern.

So the puzzle becomes: how can we respond to morally compelling cosmopolitan concerns — that the location premium is morally indefensible — without triggering political-societal rejection that makes liberalisation impossible?

This is precisely where a major part of the empirical-policy literature in economics has moved since the mid-2000s.

Michael Clemens, Lant Pritchett and co-authors have argued that the most politically feasible response to this tension is **not** to insist normatively on a binary shift from "closed borders" to "fully open borders." Instead, they propose a class of transitional institutional designs whereby *citizenship is partially unbundled*. Instead of citizenship being either full (1) or none (0), intermediate legal statuses can be created which grant *legal entry* and *labour mobility* without immediately granting the full package of political rights, entitlement to permanent residence, or family reunification (Clemens 2011; Pritchett 2006; Pritchett & Clemens 2019).

This is the move from *first-best normative justice* to *second-best political design*.

The next section will examine these proposals in detail — as an attempt to mediate, not philosophically, but institutionally — between the powerful moral case for mobility and the equally powerful political resistance to full equal membership.

9.7 Policy relevance: unbundling citizenship to lower native resistance

If the citizenship premium is real and very large — as we have seen in the previous sections — then it is unsurprising that receiving societies resist large inflows of low-skill immigrant workers. From the standpoint of voters in rich countries, admitting migrants is equivalent to *granting them access to that premium*: access to higher labour productivity, better institutional contexts, better wages, and the entire bundle of benefits that come with inclusion in an advanced political economy. Even if many voters do not articulate it in these analytical terms, this is the intuitive behavioural logic: immigration is essentially the allocation of rents (Milanović 2016; Rodrik 2018; Pritchett 2022).

In the canonical Rawlsian model the premium does not exist (Rawls 1999): "peoples" are imagined to be located in reasonably well-ordered polities, without *structural* cross-border gaps. Once the empirical magnitude of the premium is admitted, however, a crucial second step follows: *migration policy becomes distributive policy*. It is thus unsurprising that native workers in advanced economies oppose openness: they fear that they will lose — not metaphysically, but concretely — by sharing their rent.

Clemens, Pritchett and co-authors have argued that the most politically feasible response to this tension is *not* to insist on an immediate move from "closed borders" to "fully open borders," but to design transitional regimes in which "citizenship" is **partially unbundled** (Clemens 2011; Pritchett 2006; Pritchett & Clemens 2019). In other words: *not* a binary variable (full = 1 vs none = 0), but several intermediate grades.
These transitional regimes could include:

1. **Temporary legal residency with fixed time horizons** (e.g. 3–5 years)
2. **Tax discrimination or social insurance discrimination** (workers contribute more or receive fewer contributory transfers during a transitional period)
3. **No immediate access to family reunification** (family rights are rationed temporally, not denied forever)
4. **Obligation to return after period X** (unless re-authorised under a new quota cycle)

Pritchett calls this a "**vector of partial inclusion**" rather than a binary switch (Pritchett 2006, 2022). Clemens calls it "**visa policy as price discrimination over citizenship rents**" (Clemens 2011). Ruhs (2013, 2019) generalises this in terms of "**rights differentiation**": labour immigration programmes trade *numbers* against *rights* — the more rights migrants get immediately, the fewer host societies are willing to admit. The logic underlying these proposals is not normative idealism but micro-political realism: *lower the perceived distributive shock on natives*, and native resistance falls.

These second-best proposals therefore attempt to preserve the **efficiency gains** from cross-border labour mobility (Rodrik 2018; Clemens 2014) while dampening the **perceived zero-sum distributive conflict** that triggers nativist backlash. They recognise, directly, that when the premium is large and binary, voters react against immigration as if it were a direct transfer from "us" to "them." A graded system, instead, spreads the transfer over time.

In short: if we cannot erase borders tomorrow — and if voters in rich countries are unwilling to treat foreigners as full equals ex ante — then the relevant zone of political design is the space *between* "no entry" and "instant full citizenship." The task is to make mobility politically absorbable without extinguishing its developmental value. Unbundling citizenship is not utopia: it is a second-best, but a second-best that responds to both economic logic *and* political constraints.

9.8 Why a mild cosmopolitan could support more mobility

Dani Rodrik (2018) offers one of the most elegant ways to restate the moral debate on migration in empirical terms. Instead of asking, *what should justice demand across borders?*—a normative question that divides cosmopolitans and communitarians—Rodrik invites us to reframe it positively: *how much more do we have to care about our fellow citizens than we care about foreigners in order to justify keeping borders closed?*

This reframing translates moral preference into a single measurable parameter, which Rodrik calls **ϕ (phi)**. The idea is simple: ϕ represents the weight that societies implicitly place on the welfare of their own citizens compared to the welfare of non-citizens.

- If **ϕ = 1**, we are *perfect cosmopolitans*: we value the happiness of a foreigner exactly as much as that of a compatriot.
- If **ϕ is very large**, we are *complete nationalists*: only the well-being of citizens counts, while foreigners' interests are effectively ignored.

Rodrik then constructs a simple **numerical thought experiment** to see what value of ϕ would make rich countries rationally choose to keep their borders closed. Imagine that **60 million workers** from poor countries are allowed to move to rich countries—roughly a 10 percent increase in the rich countries' labour force (600 million workers)..

In this scenario, the migrants' wages rise dramatically—from **$500 per month** in their home country to **$2,000 per month** in the destination country, a fourfold increase. Meanwhile, the influx of labour slightly depresses the wages of low-skilled native workers. Drawing on empirical estimates from George Borjas, Rodrik assumes a **wage elasticity of –0.3**, meaning that a 10 percent increase in labour supply lowers wages by about 3 percent.

To evaluate this situation, Rodrik uses a **social welfare function** that adds up the well-being of both natives and migrants. Each group's utility is defined as the logarithm of its income (a common assumption in welfare economics, reflecting diminishing marginal utility). The welfare of citizens is multiplied by **ϕ**, while the welfare of migrants counts with a weight of **1**. Mathematically, the question becomes:

For what value of ϕ does this migration policy make the rich country's total welfare decrease?

Plugging in the numbers, Rodrik finds that **ϕ must be greater than 4.5** for migration to make the rich country worse off (See Box # for details of the calculation of the parameter **ϕ).**

What does this mean in plain language? It means that a rich country would need to value the welfare of its own citizens at least **4.5 times more** than that of poor foreigners in order to justify blocking migration. Conversely, if the country values foreigners at even one-quarter of the value of its citizens (ϕ = 4), then admitting migrants would *increase* national welfare.

The conclusion is striking. We do not need to be **radical cosmopolitans**—treating all people's interests equally—to justify more open borders. Even **modest cosmopolitanism**, where we think a foreigner's well-being counts for only a fraction of a citizen's, is enough to support more mobility. The reason is straightforward: the **gains for migrants** are enormous, while the **losses for natives** are small and diffuse. A migrant's income might quadruple, but a native's wage may fall by only a few percent.

Rodrik's thought experiment therefore exposes the moral arithmetic of our current migration regime. To sustain existing border barriers, we must implicitly hold an extremely high valuation of our fellow citizens relative to foreigners—so high that it verges on moral indifference to outsiders. Yet if we accepted even a mild cosmopolitan concern for non-citizens, the calculus would reverse: the case for more open migration would be overwhelming.

Rodrik's conclusion is deliberately modest but politically powerful. We do not need a world of perfect moral equality to defend greater mobility. We only need to accept that the lives of foreigners matter *somewhat*. Even under that small concession, the enormous welfare gains from migration vastly outweigh the minor domestic losses. The political question, then, is not whether open borders are utopian, but how much inequality of moral concern we are willing to tolerate to keep them closed.

9.9 Conclusion: Borders, Rents, and the Demands of Justice

This chapter has followed a straightforward but unsettling argument. Even though we speak of universal equality, a person's life chances still depend overwhelmingly on where they are born. Citizenship operates like a valuable asset that grants access to opportunities, and borders are the institutions that distribute and protect this asset. Seen this way, migration is not a disruption to the global order but the most effective way people can turn their geographical starting point into real opportunity.

The empirical patterns underline this point. Most global inequality exists between countries, not within them. Most people never move, yet those who do often multiply their earnings many times over without gaining new skills. Remittances consistently exceed foreign aid, showing that the most effective form of global redistribution still occurs through families rather than states. In this context, the current system of selective mobility—easy for the already privileged and tightly restricted for others—

appears less like neutral policy and more like a mechanism that maintains global advantage.

The moral debate mirrors this divide. Communitarian and liberal-nationalist theories argue that justice can function only within a defined political community; membership must come first, redistribution second. Cosmopolitan theories respond that if all persons have equal moral worth, this cannot stop at national borders. Preventing movement that would dramatically improve someone's life effectively defends inequalities based on accident of birth. These positions do not simply differ in emphasis—they rest on fundamentally different principles, and neither can be reduced to the other.

Political realities reflect this tension. When cosmopolitan arguments reach voters in wealthy democracies, they often provoke backlash: stricter borders, outsourcing enforcement, and limiting welfare to citizens. When states attempt tight national closure in an economically interdependent world, the result is usually workarounds: irregular migration, informal labor markets, and migrants who hold some rights but no secure status. Both patterns reveal the same underlying problem: a world organized around national membership trying to manage challenges created by global integration.

This chapter offered two approaches for making this tension more manageable. The first is institutional. By separating different elements of citizenship, states can create legal, gradual migration pathways that capture much of the economic benefit while reducing perceived shocks to social cohesion. Done carefully, this does not entrench permanent second-class status; rather, it structures inclusion in a way that politics can sustain. The second approach is diagnostic. Rodrik's ϕ thought experiment translates moral claims into numerical terms: to justify current levels of restriction, we must value citizens more than four times as much as otherwise identical non-citizens. One does not need to be a strong cosmopolitan to find that preference difficult to defend.

The broader lesson for policymakers is that the real choice is not between open borders and closed borders. It is between using migration as a tool for greater fairness or allowing inaction to reinforce inequalities through exclusion. Concrete reforms will differ across countries, but they will share key features: clear legal pathways; rights that expand over time; portability of social contributions; strong labor protections to prevent exploitation; and agreements with sending states that ensure mobility supports, rather than undermines, development.

Borders are not going away. But the distributive tasks we assign to them can evolve. If we take seriously our stated commitments—to equal moral worth, to opportunities not determined by birth, and to a vision of prosperity that extends beyond geography—then the direction of reform is evident. We can continue designing institutions around fear, or we can design them toward fairness. For millions of people, that choice determines whether their lives remain limited by where they began or opened by where they are allowed to go.

BOX # 9.2: The Ethics of Borders — A Comparative Map

Philosopher	Moral Framework	View on Borders	Core Argument	Right to Exclude?	Obligations Toward Migrants	Why Borders May/Must Open
Michael Walzer	Communitarianism	Borders essential to preserve community	Justice is “internal” to political communities; membership precedes distribution	Yes — strong right to control membership	Admit refugees in emergencies; otherwise discretion	Protect integrity of shared meanings and political community
Christopher Heath Wellman	Liberal freedom of association	Borders justified by collective autonomy	States, like individuals, may choose with whom to associate	Yes — states have a moral right to refuse	Must respect basic human rights; must not refuse refugees in peril	Freedom of association includes freedom *not* to admit
David Miller	Liberal nationalism	Borders required for self-determination & distributive justice	Nations are communities with obligations inwardly directed; solidarity needs boundaries	Yes — moderate right to set limits	Must admit refugees temporarily; some humanitarian duties	Protect national culture, trust, welfare state, duties to compatriots
Peter Singer	Global utilitarianism	Borders have no moral significance	Maximise global welfare; migration yields enormous welfare gains	No — exclusion wastes opportunities for welfare improvement	Duties to reduce suffering globally	Open borders increase global well-being dramatically
Thomas Pogge	Global egalitarian (institutional)	Borders unjustly preserve structural inequality	Rich countries co-create global poverty; blocking migration compounds harm	No — exclusion is part of an unjust global system	Duties of justice (not charity) to the poor	Migration remedies harm caused by unjust global institutions
Joseph Carens	Liberal egalitarian cosmopolitanism	Borders are morally arbitrary and unjust	Citizenship-based exclusion is like feudal privilege; freedom of movement is a basic right	No — very strong presumption against any exclusion	Must allow broad mobility; may restrict only with very strong reasons	Borders perpetuate birth-ascribed privilege; open borders follow from core liberal principles

BOX # 9.3: The Logic of Rodrik's ϕ Calculation

Rodrik's thought experiment can be expressed as a simple welfare comparison between the *status quo* (closed borders) and a *policy reform* that allows migration from poor to rich countries.

Let:

- W = social welfare of the rich country
- U_n = utility of natives
- U_m = utility of migrants
- ϕ = relative weight placed on natives' welfare compared to migrants' welfare

Then:

$$W = \phi \times U_n + 1 \times U_m$$

Utility is assumed to be **logarithmic** in income, reflecting diminishing marginal returns to money:

$$U = \ln (y)$$

where y is income.

Assumptions

Migrants allowed in: **60 million**

- This raises rich-country low-skill labour supply by **+10%** → implies about **600 million** native low-skill workers.
- Native wage before: **$3,000/month**.
  Wage elasticity to labour supply = **–0.3** → +10% supply ⇒ **–3%** wage → new native wage **$2,910**.
- Migrant wage before: **$500/month** (in origin).
  After moving: **four-fold** → **$2,000/month**.
- Social welfare in host = **ϕ × (sum of natives' utility)** + **1 × (sum of migrants' utility)**.
- Utility = **log(income)** (standard).

Compute utility changes

- Native utility change per person:
  $\Delta U_N = \ln (2910) - \ln (3000) = \ln (0.97) \approx -0.030459$.
- Migrant utility change per person:
  $\Delta U_M = \ln (2000) - \ln (500) = \ln (4) \approx 1.386294$.

Aggregate change in (host) social welfare

$$\Delta W = \phi \cdot N_N \cdot \Delta U_N \; + \; N_M \cdot \Delta U_M$$

with $N_N = 600$ m, $N_M = 60$ m.

We ask: for which ϕ does $\Delta W < 0$ (i.e., the policy *reduces* host social welfare)?

Solve for the threshold, where $\Delta W = 0$:

$$\phi = -\frac{N_M \cdot \Delta U_M}{N_N \cdot \Delta U_N} = -\frac{60 \text{ m} \times \ln 4}{600 \text{ m} \times \ln (0.97)} \approx \frac{60 \times 1.386294}{600 \times 0.030459} \approx 4.55.$$

Interpretation

- If **ϕ < 4.55** → you don't weight natives *that* much more than foreigners → the **huge** gains to 60m migrants (log(4) each) outweigh the small 3% loss spread over ~600m natives → **net welfare rises**.
- If **ϕ > 4.55** → you value natives **>4.5×** as much as foreigners → you'd judge the policy as welfare-reducing for the host.

So, you **do not** need to be a cosmopolitan (ϕ=1). Even with a **strong native preference** (a citizen counts ~4.5 foreigners), let's say being a **loose nationalist** or a **mild cosmopolitan**, the case for allowing substantially more labour mobility still goes through in this stylized calculation.

BOX # 9.4 - Rawls on Immigration: The Disappearance of the Problem in a Well-Ordered World

Rawls's discussion of immigration in *The Law of Peoples* is strikingly indirect. Unlike most contemporary treatments, which approach migration as a domain of distributive conflict between potential claimants (states, citizens, migrants), Rawls does not treat migration as a primary question of justice at all. Instead, migration is framed as an index of background injustice in the *sending state*. Immigration flows arise, on his view, only under conditions of political non-decency or institutional failure. In a Society of liberal and decent peoples, immigration would not become a serious problem, because the causes driving people to cross borders would have been eradicated.

To reach this result Rawls makes two connected claims.

First, he characterizes political society itself as the responsible "agent" charged with stewarding territorial resources and sustaining them in perpetuity. Although the boundaries of existing states may be historically arbitrary, a people, once politically organized, is entitled — indeed required — to maintain the environmental integrity and demographic sustainability of the territory for which it is responsible. Rawls thus reconceives territorial jurisdiction as analogous to a form of collective property: an asset which must be preserved, not over-exploited or offloaded onto others. Immigration is thereby linked conceptually to the stewardship function of political communities as collective property holders.

Second, Rawls ties emigration pressures to endogenous injustice. Famines, oppressive social hierarchies, religious persecution, conscription of peasants for the dynastic wars of monarchs — these are not natural facts, but political failures. Most importantly, he treats "population pressure" not as an exogenous parameter, but as a structural consequence of gender subordination: once women are fully equal, politically empowered, and educated, he claims, fertility stabilizes and population-driven migration ceases. Thus, well-ordered peoples — liberal or decent — will not export surplus people because they will have brought the underlying drivers of out-migration under institutional control. In this framework immigration is not a site of competing distributive claims, but a *symptom* of injustice elsewhere. Policy responses at the border are therefore secondary. The site of normative repair is internal reform of the sending society — not redistribution of territorial access by the receiving one. Rawls can therefore preserve strong border control implicitly, without developing a theory of admissible entry, because in the ideal — when the principles of the Law of Peoples are realized — the underlying need for movement vanishes.

The controversial consequence of this position is that the hardest questions about migration never arise. Immigration is dissolved, not adjudicated. Rawls does not construct a theory of the legitimate claims of migrants *as such*; he constructs an institutional model in which there would be no migrants. Pogge criticizes Rawls's treatment of immigration for deleting the normative problem instead of addressing it. For Pogge, immigration flows do not arise merely from domestic injustice but from an unequal global order that affluent states help to design and enforce. To then refuse entry while continuing to benefit from that unjust order is not an exercise of responsible stewardship (as Rawls claims), but an ongoing violation of negative duties. On Pogge's account, Rawls's move illegitimately relocates causal responsibility onto "sending societies," thereby shielding the structural agency of the global North in producing the very pressures that drive migration.

Chapter 10

THE CLIMATE CHOICE

10.1 Introduction

Over the last three decades, climate change has moved from the margins of public debate to the centre of global politics. International agreements—from Rio and Kyoto to Paris and Glasgow—have framed climate change as a shared challenge for humanity, and periods of technological optimism have fostered the belief that decarbonisation is not only necessary but also feasible. Yet, as this chapter is written, there are mounting signs of fragility in this consensus. Governments that once championed ambitious climate targets now hesitate or backtrack under pressure from energy price shocks, geopolitical tensions, populist movements, and powerful incumbent interests. Commitments made with great fanfare are quietly diluted, postponed, or reinterpreted. The "climate choice" is therefore not a one-off decision already taken in Paris, but an ongoing and contested process—one that can move forwards or backwards depending on political will.

This instability makes the analytical core of the climate problem more important, not less. At a purely economic level, the structure of the issue is surprisingly clear. Climate change is a textbook case of market failure: greenhouse gas emissions are an unpriced externality, and without intervention the world emits far more than is socially optimal. From this perspective, the central question is how to design policies—carbon taxes, cap-and-trade systems, regulation, public investment—that align private incentives with social welfare at minimum cost. Cost–benefit analysis, as pioneered on a grand scale by the Stern Review, provides a systematic framework for this task. It asks whether the present costs of mitigation are justified by the future benefits of avoided damage, and it forces us to confront two crucial ingredients: deep **uncertainty** about future climate impacts and the ethical problem of **discounting** future generations' wellbeing.

But the climate choice is not only an economic problem. It is also, and perhaps primarily, a question of **justice**. Even if we agree on how much global mitigation is warranted, we still face the far more contentious issue of how to share the burdens and benefits among countries. Who should cut emissions fastest? Who should finance adaptation and loss-and-damage in vulnerable regions? How should responsibilities be divided between historically high emitters and countries that remain poor but are rapidly industrialising? These questions cannot be answered by efficiency alone. They require normative principles—equal per-capita emissions, polluter pays, beneficiary pays, ability to pay, or, as John Broome proposes, a no-sacrifice regime grounded in resource-constrained efficiency and self-interest.

The political risks of our current moment arise precisely at the intersection of these two dimensions—efficiency and justice. On the one hand, the economic case for ambitious early mitigation is remarkably robust: once uncertainty, catastrophic risk, and the long-lived nature of CO_2 are properly taken into account, strong climate action appears not only morally compelling but economically prudent. On the other hand, disagreements over fairness—how to allocate costs between North and South, past and present, rich and poor—repeatedly stall or weaken international cooperation. There is a real danger that, in insisting on our preferred conception of justice without compromise, we fail to secure any binding agreement at all. In that scenario, the losers are not only future generations but also today's poorest and most vulnerable communities, who are least responsible for the problem and least able to protect themselves from its consequences.

This chapter examines the climate choice along both of these axes. In the first part, it explores the economic analysis of climate change: the role of cost–benefit reasoning, the challenges posed by uncertainty and fat-tailed risks, and the ethical controversy over how to discount the future. It then turns to the policy instruments that follow from this analysis, especially carbon pricing and the evolution of cap-and-trade systems as a central tool of global climate governance. In the second part, the chapter shifts focus to the normative question of burden-sharing. It introduces and critically assesses the main principles proposed in the literature—equal per-capita emissions, the polluter-pays principle, the beneficiary-pays principle, the ability-to-pay principle, and Broome's no-sacrifice approach—highlighting both their moral appeal and their limitations in practice.

The aim is not to defend a single, definitive solution, but to clarify the structure of the choices before us. In a world where some governments are retreating from earlier promises, it is tempting either to lower our ambitions or to insist on ideal justice at any cost. Both responses are dangerous. The task, instead, is to understand how economic efficiency and climate justice can be brought as close together as possible—how we can design institutions that are both normatively defensible and politically workable. The climate choice, in this sense, is not simply about whether to act, but about how to act in ways that are fair enough to be legitimate and feasible enough to succeed.

10.2 Are Governments Stepping Back from the Paris Agreement? Political Risks and the Future of Climate Targets

Nearly a decade after its adoption, the Paris Agreement remains the most comprehensive global framework for mitigating climate change. Its core strength—the near-universal participation of states—rests on a shared commitment to pursue efforts to limit warming to 1.5°C and to keep it "well below" 2°C. Yet recent political developments raise the question of whether governments are drifting away from the Agreement's spirit, if not its formal structure, and whether the world is still on track to meet its climate objectives.

Although no large-scale exodus from the Paris Agreement has occurred, individual governments have begun signalling withdrawal, weakening their pledges, or delaying implementation. The most visible example is the United States, which again announced its intention to leave the Agreement, citing concerns about economic competitiveness and national sovereignty. Other governments have softened their targets, pushed back net-zero timelines, or slowed regulatory reforms. While these moves do not dismantle the framework itself, they erode the collective ambition on which the Agreement depends.

This political retreat has direct implications for the likelihood of achieving global temperature goals. Current national policies and pledges fall far short of what is needed to stay within 1.5°C of warming. Even if all existing commitments were fully implemented—which is far from guaranteed—the world would still be on a trajectory toward roughly 2.3–2.7°C of warming. Scientific assessments increasingly warn that the carbon budget compatible with 1.5°C has nearly been exhausted, and several recent analyses argue that overshooting this threshold has become effectively unavoidable. In this context, any political backsliding by major emitters makes the Agreement's aspirational goals even harder to realise.

The risks to the Paris targets cannot be understood without attention to the broader political climate. Climate governance depends on long-term, coordinated action across electoral cycles, yet domestic politics in many countries remain volatile. Shifts in government can quickly change national priorities, especially where climate policy is framed as economically burdensome or culturally divisive. Powerful vested interests—particularly in fossil fuel extraction and energy-intensive industries—continue to exert influence, lobbying for weaker regulation or slower transitions. Public resistance to climate policies perceived as costly or unfair can further constrain political leaders, especially in contexts of inflation, energy crises, or geopolitical instability.

Internationally, the Paris Agreement relies on voluntary, self-determined commitments. This flexibility makes broad participation possible, but it also means that the system's credibility depends on trust: if leading emitters weaken their efforts, others may follow, fearing competitive disadvantage or questioning the value of unilateral action. In such an environment, even modest political shifts can trigger broader erosions in ambition.

In short, while the Paris Agreement has not fractured, its success is far from secure. The combination of insufficient current policies, rising global emissions, and shifting political priorities creates a real danger that climate targets will not be met. The problem is not merely technical or economic; it is fundamentally political. The future of the Paris goals will be determined as much by electoral outcomes, public opinion, and international diplomacy as by technological innovation or economic modelling. The task ahead is therefore not only to reduce emissions but to rebuild and sustai The European Union remains one of the strongest global advocates for ambitious climate action, but its position is becoming more difficult to sustain. Legally, it is still committed to climate neutrality by 2050 and to cutting emissions by at least 55% by 2030, and it continues to expand tools like carbon pricing and the new border carbon adjustment mechanism. Yet the political and economic context has shifted.

The energy crisis triggered by Russia's invasion of Ukraine created short-term setbacks and raised concerns about industrial competitiveness. Inflation, social pressures, and growing opposition from farmers' groups and right-leaning parties have led several Member States to resist or slow elements of the Green Deal. As a result, meeting the EU's own climate targets is becoming more challenging.

Internationally, the EU still plays a stabilising role, but its influence is weakened by competition with the U.S. Inflation Reduction Act and by calls from developing countries for more climate finance.

Overall, the EU remains a key climate leader, but political fragmentation and economic pressures mean its leadership is less assured than it once seemed.

n the political commitment necessary for meaningful global climate action.

If governments today appear increasingly willing to delay or dilute their climate commitments, this trend must be understood against the backdrop of a widening gap between political incentives and economic fundamentals. The physical processes driving climate change have not paused, nor have the projected damages become less severe. What has changed is the political environment: rising geopolitical conflict, energy-price volatility, and populist pressure have made climate ambition more fragile. Precisely because political will is wavering, it becomes essential to revisit the economic logic that initially underpinned global climate cooperation. To appreciate why mitigation remains sound policy—indeed, why it becomes more urgent as delays accumulate—we now turn to the framework of cost–benefit analysis. This analytical lens, especially its treatment of uncertainty and discounting, clarifies the scale of the long-term risks and the rationale for acting before those risks become irreversible.

10.3 Cost–Benefit Analysis of Climate Change: Uncertainty, and Discounting

The **Stern Review (2006)** was the first major and systematic application of cost–benefit analysis to the climate problem. It framed climate change not merely as an environmental threat, but as an economic choice involving present costs and future benefits. Its central claim—that strong early mitigation is economically justified—rested explicitly on two analytical pillars that remain crucial in any serious assessment: **(a) incorporating uncertainty and the behavioral effects associated with risk**, and **(b) determining how future outcomes should be discounted**. These two elements shape the evaluation of virtually every climate policy.

A rigorous cost–benefit framework begins with the physical foundations of the climate system. The Earth's temperature is governed by the balance between incoming solar radiation and outgoing infra-red radiation. While nitrogen and oxygen—the dominant atmospheric gases—are transparent to both, a small set of "trace gases," notably carbon dioxide, methane, and water vapor, absorbs infra-red radiation and traps heat. Since the Industrial Revolution, human activity has sharply increased the concentration of these gases, raising atmospheric CO_2 from roughly **280 ppm** pre-industrially to around **430 ppm CO_2-equivalent** today. Depending on the policy action that government will take in the next decades concentrations may range between 500 and **550 ppm by 2050**, a level not seen for several million years,

Climate models associate these concentrations with substantial warming. Stabilizing at 550 ppm it is likely to generate at least **2°C** of global average temperature increase. Under business-as-usual, there is a significant probability—on the order of **50%**—that warming could exceed **5°C** by 2100. The full consequences of climate change are not well-know, but many effects are quite foreseeable:

- Bad implications for agriculture especially in heavily populated tropical regions.
- Rise in sea-level which will wipe out the small island countries and coastal towns forcing millions of people to migrate.
- The rise could be catastrophically rather than gradual if Greenland and West Antarctic melt and collapse.
- Reversing of the Gulf Stream could cause climate in Europe to resemble that of Greenland.
- Tropical storms would become more severe causing famine and drought.
- Glaciers will disappear and with them valuable water supply.
- The poorest economies will be the most affected, thus increasing inequality and extreme poverty.

The loss of welfare of the human beings and also many sentient beings which in its calculation should also include deaths and life shortenings could be immense.

Two structural features of climate change complicate the use of cost–benefit analysis. **First**, greenhouse gases are **extremely persistent**: once emitted, CO_2 remains in the atmosphere for centuries. Unlike pollutants that dissipate after emissions cease, carbon dioxide creates long-lasting climatic effects.
Second, climate change is a **fully global externality**. Greenhouse gases mix in the atmosphere within days, so the costs of an emission are distributed globally, regardless of where it originates. This makes climate change a paradigmatic *global public bad* and limits the effectiveness of unilateral mitigation.

Within this framework, cost–benefit analysis compares the **benefits of mitigation**—the avoided damages of future climate change—with the **costs of mitigation**, namely reduced current and future consumption due to shifts toward cleaner fuels, improved efficiency, changes in behavior, or reduced output. These measures can be represented as imposing a **permanent fractional cost on global GDP**.

To evaluate whether mitigation is justified, these future benefits must be brought into present value terms. This requires a **social discount rate**, typically decomposed as:

$$\delta = \rho + g\eta,$$

where

- **ρ** is the pure rate of time preference,

- **g** is the projected growth rate of consumption, and
- **η** captures the declining marginal social value of consumption in wealthier societies.

The term $g\eta$ is widely accepted in form, though its magnitude is debated. Since future generations will presumably enjoy a higher standard of living than present generations if economic growth will not halt, it is morally acceptable to discount the avoid damages of future climate change. By contrast, the choice of **ρ** is deeply controversial. Many philosophers and some economists argue for **near-zero pure time preference**, rejecting the idea that future wellbeing should be devalued simply because it occurs later. Others argue that a positive ρ is necessary to reflect observed savings behaviour and individual preferences. The value chosen for ρ heavily influences assessments of climate policy.

The economist Tyler Cowen and the philosopher Derek Parfit argue that using a positive social discount rate to devalue future wellbeing is morally indefensible and leads to systematically underinvesting in policies that benefit future generations.
They claim that discounting future welfare *simply because it occurs later* is ethically arbitrary and equivalent to discriminating against future people. Instead, only changes in wealth, technological progress, and uncertainty should affect how we value the future—not time itself. Discounting future utilities implies that we may knowingly permit catastrophic harm to future generations for small present benefits, a conclusion they reject as morally unacceptable. Therefore, they argue for **zero pure time preference** and for evaluating policies based on their real effects on future people's wellbeing, not their temporal location (See Box # 1) for a discussion about discounting and intergenerational ethics.

Beyond discounting, cost–benefit analysis must incorporate **uncertainty and risk aversion**.

A central limitation of simple cost–benefit analysis is that it often evaluates climate damages using their **expected value**, as if only the average outcome matters. Yet climate change is characterized not merely by uncertain future losses, but by the possibility of **extreme, potentially catastrophic events**. Once we take seriously how individuals and societies actually evaluate uncertain outcomes—given risk aversion—this uncertainty profoundly strengthens the case for early mitigation. A simple example can illustrate better this concept. Imagine you can either:

- Pay €1,000 now to protect your house from floods, or
- Pay nothing, but face a 0,5% chance of a flood that would destroy your home and cost €200,000**.**

Even though the *expected* average cost of the risky option is also €1,000 (0,5% X €200,000), and so people should be indifferent between to be insured or not to be insured, most people would choose to pay the certain amount. The small chance of a huge disaster feels much worse than the arithmetic suggests.

Climate change is similar. Most forecasts predict serious but manageable harm, but there is also a small—yet real—chance of catastrophic outcomes: rapid ice-sheet melt, major ocean circulation changes, runaway warming, or massive ecosystem collapse. These "tail risks" matter a lot because their consequences are enormous.

When we take these extreme possibilities seriously, the cost of doing nothing becomes much higher than simple averages indicate. This makes climate mitigation extremely valuable. Cutting emissions is like paying the house-protection cost: it reduces the chance of disaster. Even modest reductions in catastrophic climate risks provide huge

As Martin Weitzman (2009) and others have stressed, conventional discounting formulas and expected-value calculations underestimate the true danger because they cannot capture the weight that informed, risk-averse agents place on catastrophic tail events.

This reasoning greatly increases the value of mitigation. By reducing the probability of extreme warming—even modestly—strong climate policy delivers welfare gains far exceeding what traditional models identify. Mitigation becomes not only an investment in expected future wellbeing but also a form of **insurance** against global catastrophe. In this insurance framework, the premium—current mitigation costs—is small relative to the value of avoiding even a small chance of irreversible, civilization-scale harm (See box # 2 for an evaluation or existential risk).

Thus, when cost–benefit analysis properly incorporates risk aversion and the fat-tailed nature of climate uncertainty, the justification for ambitious climate action becomes significantly stronger. Far from weakening the economic argument for early mitigation, uncertainty transforms climate policy into one of the most rational and prudent collective investments available.

A simple example shows why the basic conclusion of cost–benefit analysis is so robust. Imagine that, if we do nothing, the world economy will suffer very large climate damages over the next two centuries—equivalent to losing roughly one-fifth of future global income. In economic terms, this is like slowing down long-run growth: instead of the world economy growing at about 1.3% per year, it would grow at only 1.2%. Over many decades, this small difference compounds into enormous losses.

Now compare this with the cost of preventing those damages. Stabilizing greenhouse-gas concentrations at safer levels might require the world to spend around 1% of its income every year on cleaner technologies, better efficiency, and reduced emissions. This is a noticeable but manageable cost—especially when weighed against the much larger long-term losses that would result from inaction.

The key question is therefore simple: Is it worth paying a modest cost today to avoid very large losses in the future?
When economists translate these numbers into present-value terms, the answer is "yes" under almost any reasonable assumption. Even people who place a **very low value on the future**—who strongly prefer present benefits over future wellbeing—

would still judge early mitigation to be worthwhile unless they adopt an implausibly high rate of time preference. Only if we assumed that future generations matter extraordinarily little (far less than any ethical or empirical view would accept) would mitigation fail a cost-benefit test.

In other words, the conclusion is extremely robust: across a wide range of ways of valuing the future, including views that are ethically questionable, the long-run benefits of climate action substantially exceed the costs. Far from being sensitive to philosophical assumptions, the economic case for mitigation remains strong even under discounting choices that heavily favor the present at the expense of future generations.

This does not imply that all uncertainties are resolved. Damage estimates—particularly for high-temperature outcomes—remain contested, and mitigation costs vary with technological progress. Nonetheless, the structure of the problem—**extremely persistent emissions, non-negligible probabilities of catastrophic outcomes, and the compounding effects of small changes in growth rates over centuries**—means that even conservative economic assumptions support substantial early mitigation.

In sum, when cost–benefit analysis is applied with explicit attention to **uncertainty** and **the ethical and empirical challenges of discounting**, it strongly endorses ambitious near-term mitigation. Preventing dangerous climate change emerges not only as a moral imperative, but also as sound, forward-looking economic policy.

BOX # 10.1: Discounting and Intergenerational Ethics: A Comparative Discussion

The debate over how to discount the future sits at the centre of climate economics, and the choice of a social discount rate is one of the most consequential normative decisions in policy analysis. Although the formal apparatus traces back to Frank Ramsey's early twentieth-century work on optimal savings, contemporary interpretations diverge sharply, producing different recommendations for climate action. The controversy reflects deeper philosophical disagreements about the moral standing of future generations, the weight of uncertainty, and the role of economic growth. The Stern Review, the critiques of Tyler Cowen and Derek Parfit, the refinements proposed by Partha Dasgupta, and the cost-effectiveness concerns raised by John Broome all illustrate different facets of a long-standing debate.

1. Ramsey's Framework and the Origins of the Discount Rate

Ramsey's (1928) optimal growth model provides the canonical starting point. In Ramsey's formula, the social discount rate decomposes into:

$$\delta = \rho + g\eta,$$

where

- **ρ** reflects pure time preference (how much less we value the mere fact that an event occurs later),
- **g** is the expected growth rate of consumption, and
- **η** captures the elasticity of marginal utility—the declining value of an additional unit of consumption in wealthier societies.

Ramsey himself argued that a positive **ρ** is "ethically indefensible" and should be set to zero. Future wellbeing should not be discounted simply because it occurs in the future. Any discounting of future utility should stem from the declining marginal utility of consumption (η) and the expectation that future people will generally be richer (g), not from an arbitrary preference for the present over the future.

Yet, despite Ramsey's original stance, later economists incorporated a positive ρ to reflect observed savings behaviour, impatience, and political feasibility. This divergence between Ramsey's ethical claim ($\rho = 0$) and empirical practice ($\rho > 0$) marks the beginning of the modern dispute.

2. The Stern Review: A Revival of Ramsey's Ethical Purity

The **Stern Review (2006)** revived Ramsey's ethical intuition by adopting a near-zero pure rate of time preference ($\rho \approx 0.1\%$). Stern argued that future generations should not be treated as morally inferior merely because they live later. Like Ramsey, Stern grounded discounting in the second term of the formula: **future consumption will likely be higher**, so the marginal social value of an additional unit of consumption will be lower.

Stern's choice of a low discount rate dramatically raises the weight of future climate damages in present-value terms. As a result, his analysis concludes that the benefits of aggressive mitigation far exceed its costs. In effect, Stern treats future people as ethical equals and emphasises **the asymmetry created by the persistence of emissions** and the non-negligible probability of catastrophic warming.

3. Cowen & Parfit: The Moral Case Against Positive Time Preference

In "Against the Social Discount Rate," **Tyler Cowen and Derek Parfit** sharpen Stern's ethical argument into a direct philosophical critique of positive time preference. They argue that discounting future utility because it occurs later is equivalent to a morally arbitrary prejudice — "temporal discounting" akin to discrimination based on location, identity, or race. Future persons, once they exist, have the same moral status as present persons; therefore, **a positive pure time preference cannot be justified**.

For Cowen and Parfit, permissible discounting may occur only through the diminishing marginal utility term (η) and expected growth (g), never through ρ. A positive ρ is not simply unmotivated; it is a violation of basic ethical symmetry. In the context of climate change, they argue, positive time discounting effectively allows present generations to inflict catastrophic harm on future people for relatively minor present benefits—a conclusion they deem indefensible.

4. Dasgupta: Ethical Constraints, Empirical Realities, and Social Welfare

Partha Dasgupta occupies an intermediate position. He agrees with Ramsey, Stern, and Parfit that pure time preference should ideally be zero. However, Dasgupta emphasises two additional considerations:

1. **Intergenerational equity must be constrained by empirical realities**: resource scarcity, population growth, and environmental degradation may reduce future welfare despite rising incomes.
2. **Observed savings behaviour and political feasibility matter**: societies do not behave as if $\rho = 0$, and climate policy must operate within this behavioural framework.

Dasgupta thus argues for an approach that respects the ethical case against discounting future utility yet acknowledges that the parameters g and η must be estimated from real economies, not chosen freely. For Dasgupta, the central question becomes whether **actual growth prospects** justify substantial discounting or whether climate change will erode future consumption so severely that g approaches zero or becomes negative—implying much lower discounting in practice.

5. Broome: No-Sacrifice Climate Policy and the Limits of Discounting

John Broome approaches discounting from a different angle. While he accepts that the discounting debate has ethical significance, he argues that disagreements over ρ have obscured a deeper point: **climate mitigation can in principle be designed so that present generations bear no net sacrifice at all**. Through appropriate fiscal transfers—such as raising carbon prices while reallocating revenues—governments can offset present burdens, making climate policy compatible with the self-interest of current citizens.

Broome's "no-sacrifice principle" reframes the issue: if climate policy can be implemented without reducing present welfare, then arguments about the relative weight of future utility become less contentious. In this sense, Broome shifts attention from pure discounting to **institutional design**, arguing that moral and economic considerations can be reconciled if policies are structured efficiently.

Despite their differences, all four perspectives converge on a crucial point: **the long-lived nature of carbon emissions and the possibility of catastrophic future damages render standard high discount rates inappropriate for climate policy**. Even modest changes in growth or catastrophic risk magnify the welfare losses of inaction, strengthening the economic case for early mitigation.

BOX # 10.2: The Astronomical Stakes: Understanding the Magnitude of Existential Catastrophe

One of the most profound contributions of Nick Bostrom's work (*Bostrom, 2002 & 2013*) on existential risk is his insistence that we reconsider what "serious risk" truly means when applied to the future of humanity. Much moral reasoning tends to focus on immediate harms—loss of life, suffering, economic devastation, or social breakdown. Yet Bostrom argues that when we extend our evaluative horizon to include the **entire future of human (and post-human) civilisation**, the scale of possible loss becomes unlike anything captured by conventional concepts of tragedy or disaster.

This perspective becomes clear when Bostrom examines the **magnitude of expected loss** from existential catastrophes. He begins by noting that many people attempt to grasp the stakes by extrapolating from the worst disasters in recorded history: the world wars, the Holocaust, the Spanish flu pandemic. These events, each catastrophic in human terms, have shaped cultures, borders, and moral consciousness. Yet, as Bostrom shows, even these calamities barely register on a graph of global population. In terms of the long arc of human existence, they have not fundamentally interrupted civilisation's technological, demographic, or economic trajectory.

This observation leads to a striking conclusion: **even the gravest historical tragedies fall orders of magnitude short of existential catastrophe**. The reason is simple but often overlooked. Historical catastrophes harmed humanity deeply, but temporarily; we recovered, repopulated, rebuilt, and continued the project of civilisation. **Existential catastrophes differ not in degree but in kind**: they destroy not only the present but the future.

To make the point vivid, Bostrom invokes Derek Parfit's famous thought experiment. Parfit asks us to compare three outcomes: (1) peace, (2) a nuclear war killing 99% of humanity, and (3) a nuclear war killing 100%. While most people intuitively treat (1) versus (2) as the larger moral gap, Parfit insists the real discontinuity lies between (2) versus (3). The death of 99% of humanity would be a catastrophe beyond imagination, yet **survivors could eventually rebuild**; the chain of generations would continue; the possibility of future flourishing would remain open. But the death of 100%—extinction—ends all future possibilities. It eliminates not only the value held in existing lives but also **the entire potential of our species**, stretching perhaps billions of years into the future and encompassing countless billions or trillions of future individuals.

Bostrom pushes this idea further by quantifying, at least conceptually, what this loss entails. Even if we consider the most conservative scenario—humanity never leaving Earth, never achieving radical technological progress, never colonising space—the total possible number of future human lives is enormous: at least **10^{16} lifetimes** on Earth alone over the planet's remaining habitability. Once we expand the frame to include technologically mature civilisations, space colonisation, and digital minds, the potential value becomes literally astronomical. Estimates rise to **10^{34}**biological life-years, **10^{54}** subjective life-years of artificial minds, or even higher depending on the assumptions.

Once this shift in perspective is made, the ethical implications become striking. If the future could contain **astronomically vast amounts of value**, then the expected loss from even a tiny probability of existential catastrophe far outweighs the moral importance of nearly any ordinary action. As Bostrom puts it, reducing existential risk by a single millionth of a percentage point could have an expected value greater than saving a million lives today. A reduction by a billionth of a percentage point may outperform any philanthropic or humanitarian effort we can currently imagine.

This is not a claim that present lives do not matter. Rather, it is an argument about the **structure of moral stakes** when future generations are included. Because existential risks threaten the entire long-term future, they become **uniquely morally urgent**. Their prevention dominates much of long-term moral reasoning, at least within any framework that values future life—even minimally.

Importantly, Bostrom stresses that existential catastrophes need not involve literal extinction. “Permanent stagnation,” “flawed realisation,” or “subsequent ruination” could also destroy most of the future’s potential value while leaving some humans (or human-like entities) alive. What matters is not survival per se but the **loss of humanity’s long-term potential**—the closing of the window through which civilisation might pursue knowledge, flourishing, justice, art, or moral improvement on scales vastly larger than today.

This reorientation—from disasters measured by present casualties to catastrophes measured by the destruction of future value—is what makes Bostrom’s claim so radical. It shifts our moral attention from familiar tragedies to unfamiliar but vastly more consequential possibilities. The magnitude of existential risk, once understood, forces us to see that the future is not merely an extension of the present; it is a realm of incredible moral importance. An existential catastrophe is not simply “much worse” than historical catastrophes—it is a loss so immense that the comparison becomes meaningless.

The central insight is unsettling but transformative: what hangs in the balance is not just our species, but the entire story that humanity has yet to write.

10.4 Carbon Pricing, Market Failure, and the Evolution of Cap-and-Trade

Figure 10.1 illustrates the basic economic logic behind climate change as a market failure. The horizontal axis measures **emissions**, while the vertical axis measures the **cost per ton of CO_2**. The **marginal private cost** (MPC) curve represents the cost borne directly by firms when they emit. In contrast, the **marginal social cost** (MSC) curve incorporates the **external harm** imposed on others through climate damages—future sea-level rise, heatwaves, crop losses, ocean acidification, or the heightened risk of catastrophic events.

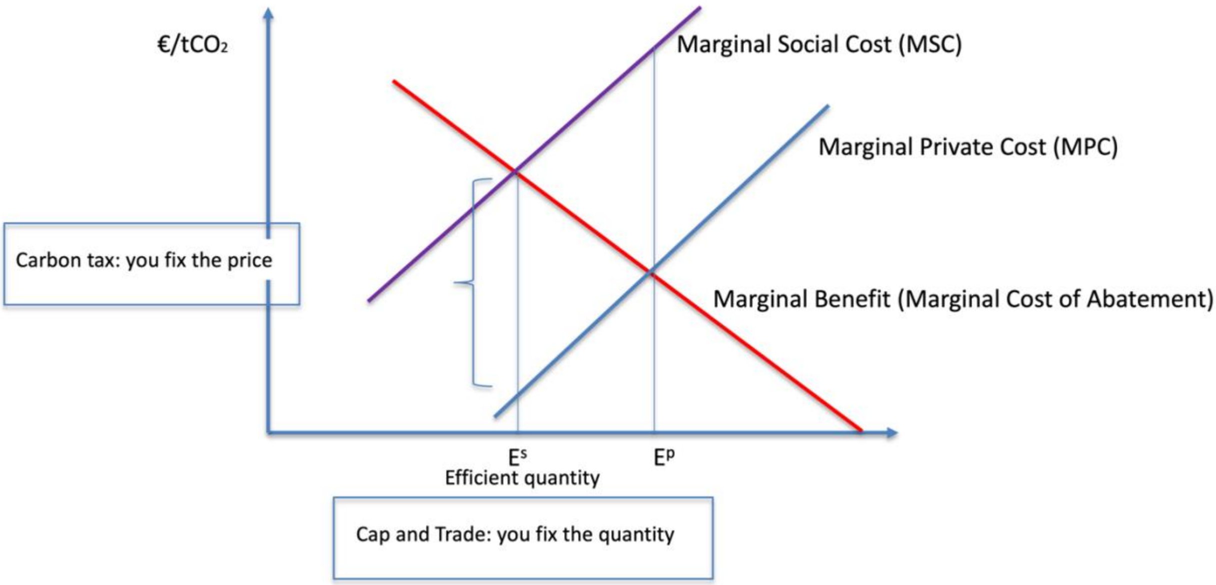


Fig.10.1

The gap between these two curves reflects the **unpriced externality** at the heart of climate change. Because emitters pay only MPC rather than MSC, they choose an emissions level E^p where their private marginal cost intersects the **marginal abatement (or marginal benefit) curve**. That curve represents the declining marginal benefit from emitting (equivalently, the rising marginal cost of abatement): at low levels of abatement, emissions are cheap to reduce; at high levels, they become increasingly costly.

The socially efficient level of emissions E^s is where the marginal abatement curve intersects **MSC**, not MPC. Since $E^p > E^s$, the market overproduces emissions, generating the shaded triangle of **deadweight loss**—a welfare loss society bears due to the divergence between private and social incentives. The solution is conceptually simple: **force emitters to internalise the externality** by charging them a price equal to the external damage their emissions cause.

The Rationale for a Carbon Price

If each firm paid a price per ton of CO_2 equivalent to the **social cost of carbon (SCC)**—the monetised value of climate damages—then the MPC curve would shift upward until it coincides with MSC. At this point:

- firms' private incentives align with societal interests,
- emissions automatically fall to E^s,
- the deadweight loss disappears,
- the economy moves to an efficient allocation.

This is why **economists overwhelmingly support carbon pricing** (Pigou 1920; Stiglitz & Stern 2017). In principle, a carbon tax equal to the SCC represents the most direct and elegant instrument.

In practice, however, determining and imposing an SCC-based carbon tax is politically and technically difficult.

Technical uncertainty

There is **no consensus** on the correct SCC because it depends on:

- the **discount rate** (Nordhaus 2008; Stern 2007; Arrow et al. 2013),
- future climate damages and tipping points,
- socioeconomic trajectories,
- probability distributions including low-probability high-damage outcomes.

Estimates range from **$30 to over $300 per ton**, with even higher values under ethical low-discount-rate assumptions.

Political resistance

Carbon taxes are visible and often regressive without compensation. Governments worry about:

- **voter backlash** (France's "gilets jaunes" movement),
- competitiveness concerns in energy-intensive sectors,
- distributional effects on low-income households,
- lobbying by powerful industrial groups.

For these reasons, many governments prefer alternatives that make the price signal less politically salient.

Partly to avoid the political difficulties of explicit taxation, the post-Kyoto climate architecture increasingly turned toward **cap-and-trade systems**.

Under a cap-and-trade scheme:

1. **A cap** sets a maximum total amount of emissions for covered sectors.

2. **Permits** (allowances) equal to the cap are issued—either auctioned or distributed freely.
3. Firms must hold one permit per ton of CO_2 emitted.
4. A **market for permits** emerges:
    - Firms with low abatement costs reduce emissions and sell permits.
    - Firms with high abatement costs buy permits instead.
5. The **permit price** becomes the economy-wide carbon price.

Instead of ensuring **price certainty** (as a tax does), cap-and-trade ensures **quantity certainty**: total emissions fall to the level of the cap.

The Global Spread of Cap-and-Trade

Cap-and-trade systems have gained remarkable traction in climate governance over the past two decades, largely because they combine economic efficiency with political pragmatism. Unlike a carbon tax—which imposes an explicit and often highly visible price on carbon—cap-and-trade embeds the cost of emissions within a permit market. Firms experience this cost indirectly, through the price of allowances, rather than as a labelled "tax," a distinction that often softens political opposition. Governments can further ease resistance by distributing a portion of permits for free during transitional periods, thereby protecting energy-intensive and trade-exposed sectors while still preserving the underlying price signal.

Beyond political considerations, cap-and-trade offers several functional advantages. It guarantees environmental performance because the total level of emissions is fixed by the cap, ensuring that the policy delivers its intended reduction. At the same time, it lowers the economic cost of mitigation by allowing firms to trade allowances, ensuring that emissions reductions occur where they are cheapest. Finally, caps can be tightened progressively over time, providing a clear and credible long-term signal to investors about the transition to a low-carbon economy.

The State of Cap-and-Trade Today

From its origins in the Kyoto Protocol era, cap-and-trade has evolved into one of the most significant tools of global climate policy. Nowhere is this more evident than in the **European Union Emissions Trading System (EU ETS)**, launched in 2005 and now the world's most mature carbon market. Covering power generation, heavy industry, and intra-European aviation, the EU ETS has undergone substantial reforms that have transformed its effectiveness. After years of oversupply and low prices—often below €10 per ton before 2013—the introduction of the **Market Stability Reserve** and the tightening of the 2030 emissions cap under the **European Green Deal** have pushed permit prices into the €60–100 range, significantly altering investment incentives and accelerating coal phase-outs across the continent.

China has emerged as another major arena of cap-and-trade innovation. The **Chinese national ETS**, launched in 2021, is now the world's largest by volume of emissions covered. Initially limited to the power sector, it is expected to expand gradually to include heavy industry, cement, and steel. Although still developing in terms of enforcement, transparency, and allowance allocation, it is set to become the central instrument of China's decarbonisation strategy.

In **North America**, California's cap-and-trade system—linked with Québec—has developed into a stable and increasingly stringent market covering multiple sectors of the economy. Meanwhile, the **Regional Greenhouse Gas Initiative (RGGI)**, covering power-sector emissions in several eastern U.S. states, has contributed to significant emissions reductions while generating substantial revenue for energy-efficiency programs and climate investments.

Similar systems operate in **South Korea, New Zealand, and the United Kingdom**, each reflecting local political and economic contexts but sharing the same fundamental design principles. At the same time, many developing countries are piloting carbon markets under **Article 6 of the Paris Agreement**, exploring linked markets, crediting mechanisms, and cooperative mitigation approaches.

Taken together, these systems now regulate **nearly one-quarter of global greenhouse gas emissions**, a dramatic expansion from almost zero in 2005. This growth reflects not only the increasing maturity of carbon markets but also the recognition that market-based instruments can deliver emissions reductions at scale and at comparatively low cost.

The intellectual rationale behind cap-and-trade is the same now as when A.C. Pigou first articulated the theory of externalities: individuals and firms should bear the **full social cost** of their actions. Carbon pricing—whether implemented through a carbon tax or a cap-and-trade system—remains the most efficient means of correcting the market failure at the heart of climate change. By aligning private incentives with social welfare, carbon pricing moves emissions from the excessive market equilibrium E^p toward the efficient level E^s. Although uncertainty, distributional concerns, and political resistance complicate policy implementation, the accelerating expansion and tightening of carbon markets across the world demonstrate that carbon pricing has become a central pillar of climate governance. Despite ongoing challenges, the trajectory is unmistakable: the cost of emitting carbon is steadily rising, and global policy is gradually converging toward a system in which emissions are constrained by economic signals that reflect their true social harm.

From Economic Efficiency to Global Justice: The Challenge of Distributing Climate Burdens

The economic diagnosis of climate change is, in many respects, disarmingly simple. As the previous sections have shown, carbon emissions represent a classic case of market failure: emitters face only their **marginal private cost**, while the true **marginal social cost**—the harm imposed on others—is far higher. Instruments such as carbon taxes or cap-and-trade systems can correct this misalignment by attaching a price to carbon that reflects its external damage. From the perspective of welfare economics, the logic is clear, and the technical tools are well established.

Yet the apparent simplicity of the economic problem stands in stark contrast to the profound difficulty of the **moral and political problem** that follows. If global emissions must fall—if the total quantity of carbon allowed into the atmosphere must be capped—then the next question becomes unavoidable: **how should the burden of reducing emissions be distributed among countries?** This issue has been the central stumbling block in international climate negotiations from Kyoto to Paris and beyond.

All countries agree, in principle, that mitigating climate change is necessary. But mitigation carries **opportunity costs**: reduced short-term economic growth, foregone industrial expansion, and the need for costly technological transitions. For developing countries, these costs may threaten poverty reduction or energy access. For developed countries, they implicate entrenched infrastructure and industrial competitiveness. The challenge, then, is not merely technical but deeply normative: **how should these collective costs be shared fairly?**

The duty to limit climate change is ultimately owed to **future generations**, who will bear the consequences of our inaction but have no voice in today's decisions. This intergenerational dimension gives the climate problem a unique moral structure. We are asked to impose costs on ourselves now—costs that may fall unevenly across nations—for the sake of people who do not yet exist. But debates about justice among present-day nations frequently block progress. Historical emitters argue that developing countries must now contribute more, while developing nations invoke historical responsibility and the right to development. Emerging economies insist that per-capita equality is essential, whereas wealthy nations stress feasibility and capacity.

Beneath these disagreements lies a deeper tension between **justice** and **welfare**. Insisting too rigidly on a particular principle of fairness—whether it be equal per-capita emissions, polluter pays, beneficiary pays, or ability to pay—may make global consensus impossible. Yet without an agreement, global emissions continue to rise, and the real victims are not negotiating states but **future generations**, along with the world's poorest and most vulnerable communities today. In this sense, excessive insistence on ideal justice can itself become unjust, because failure to reach agreement may produce far greater harm than accepting a second-best but workable compromise.

The following sections explore the main principles proposed for allocating climate burdens—**equal per-capita emissions, the polluter-pays principle, the beneficiary-pays principle, the ability-to-pay principle, and Broome's no-sacrifice principle**. Each offers a distinctive account of fairness; each embodies powerful moral intuitions; and each faces serious objections. Understanding these principles is essential for grasping why the global climate regime has struggled to find an enduring, legitimate basis for cooperation. Only by examining both the strengths and limits of these normative approaches can we appreciate the ethical stakes of climate policy—and the urgent need to reconcile theories of justice with the practical necessity of reaching agreements.

10.5 Equal Per-Capita Emissions and the Distribution of Atmospheric Rights

One influential proposal in the climate justice literature—most prominently associated with **Peter Singer** and other cosmopolitan theorists—is the allocation of **equal per-capita emission rights**. Singer's argument is grounded in the claim that all persons possess equal moral worth, and therefore each individual has an equal claim to the atmosphere as a shared global common. On this view, fairness requires dividing the remaining global carbon budget **equally among all human beings**, regardless of nationality. As Singer puts it, if we accept the moral equality of persons, "there is no non-arbitrary basis for allocating more of a shared resource to one person than to another."

The appeal of this principle becomes especially clear when we translate it into numerical terms. Suppose the remaining 1.5°C-consistent global carbon budget is roughly **400 $GtCO_2$**. With a world population of about **8 billion**, each person would be entitled to approximately **50 tonnes** of lifetime emissions over the coming decades, or roughly **0.6 tonnes per person per year** if spread evenly to mid-century. The scale of existing global inequality in emissions becomes immediately apparent: the average American emits approximately **15 tonnes annually**, the average Chinese citizen **7–8 tonnes**, while many countries in Sub-Saharan Africa emit less than **0.5 tonnes** per capita.

These disparities highlight the practical implications for developing countries. States such as **India**, with per-capita emissions around **2 tonnes**, would gain significant developmental space under an equal-per-capita regime. Many low-income African states would receive quotas far exceeding their current emissions and could, under a global trading scheme, **sell surplus allowances** to wealthier states. This would function as a substantial transfer of wealth from high-emitting to low-emitting nations, supporting electrification, industrialisation, and adaptation. In this respect, the equal-per-capita proposal converges with what **Henry Shue** identifies as "subsistence emissions": emissions needed for securing basic human needs. Many developing countries have not yet met even these basic thresholds, and equal shares would protect their right to do so.

China's position is more complex. Although China's aggregate emissions lead the world, its **per-capita** emissions of around 7.5 tonnes are still far below those of the United States. Under an equal-per-capita allocation, China would need a dramatic reduction to reach the global entitlement of roughly 0.6 tonnes. Yet China can legitimately argue—echoing Shue's emphasis on historical responsibility—that its high emissions over the last two decades have supported the greatest poverty reduction in human history, lifting more than 800 million people out of extreme deprivation. Equal-per-capita allocation, critics note, fails to integrate this developmental achievement, nor does it adequately incorporate the fact that China's cumulative emissions per person remain far lower than those of Western industrialised nations.

For the **United States**, the implications are stark. With per-capita emissions exceeding 15 tonnes and cumulative historical emissions surpassing **400 billion tonnes**—roughly one-quarter of all emissions since 1750—the United States has already vastly overused the atmospheric commons. Under strict per-capita equality, it would need to cut emissions by around **95%** and purchase significant emissions rights from developing countries. Politically, this would entail massive financial transfers outward and a wholesale reconfiguration of economic infrastructure inward. Historically, such obligations have proven unacceptable: the United States rejected the Kyoto Protocol in part because it did not impose binding commitments on China and India, yet an equal-per-capita system would impose even *greater* obligations on the United States.

These numerical projections illuminate both the moral clarity and the practical fragility of the equal-per-capita proposal. Its foundational principle of equal moral worth is compelling, especially when measured against the profound inequalities of the existing global emissions landscape. Moreover, as Singer and other advocates emphasize, its simplicity provides a powerful moral benchmark: if we truly believe in moral equality, equal access to a shared atmospheric resource follows naturally.

Yet substantial critiques have been raised. **Henry Shue** argues that equal-per-capita rights ignore both **historical responsibility** and the distinction between subsistence emissions (necessary for basic needs) and luxury emissions (which contribute little to well-being). Equal allocation may thus reproduce rather than correct structural inequities. Moreover, Broome emphasizes that an equal-per-capita system would require high-emitting countries to make massive sacrifices—sacrifices that governments have shown themselves unwilling to make. For Broome, this is why equal-per-capita allocation, however attractive morally, is **politically unworkable.**

To clarify these contrasts, the table below summarizes how the equal-per-capita allocation would apply to selected countries.

Table 1. Equal Per-Capita Emission Rights: Approximate Implications

Country / Region	Current Per-Capita Emissions (tCO_2/year)	Equal Per-Capita Entitlement (tCO_2/year)	Implication
United States	~15	~0.6	Must reduce emissions by ~95%; large net importer of emission rights; major financial obligations.
China	7–8	~0.6	Must reduce emissions by ~90%; transitional burden high, though cumulative per-capita emissions remain lower than U.S. or EU.
European Union	~6	~0.6	Must reduce by ~90%; similar structural transition to the U.S., though smaller population implies lower total transfers.
India	~2	~0.6	Moderate reductions needed; may receive surplus rights depending on population growth and development needs.
Sub-Saharan Africa (typical)	<0.5	~0.6	Gains surplus rights; potential major beneficiary in global trading system; supports energy access and development.
Global Average	~4.7	~0.6	Requires rapid global contraction; developing countries gain space, industrialised countries face massive reductions.

Sources: Global Carbon Project; IPCC AR6; World Bank population data. Numerical values approximate.

Most scholars today treat equal-per-capita allocation not as the final answer, but as a **baseline moral principle** to be combined with others—historical responsibility (Shue), intergenerational resource fairness (Broome), or hybrid frameworks such as "contraction and convergence" or "Greenhouse Development Rights." Nevertheless, its importance is ongoing: few other proposals reveal so clearly the moral stakes of climate justice and the extent of existing global inequality. As an ideal benchmark, it forces us to confront the uncomfortable reality that the atmosphere has been appropriated by a minority of the world's population, and that any just climate regime must reckon with that legacy.

10.6 The Polluter Pays Principle

Among the principles proposed to allocate the burdens of global climate mitigation, the **polluter pays principle (PPP)** occupies a central and longstanding position. Rooted in ideas of responsibility, fairness, and corrective justice, PPP holds that those who cause harmful emissions should bear the costs of addressing them. In its simplest formulation, PPP requires that "the agent who created the problem should fix it." This intuition underpins much of environmental law, informs the OECD's 1972 guidelines, and is reflected in Principle 16 of the Rio Declaration. In the climate context, it is most influentially defended by **Henry Shue**, **Simon Caney**, and **Edward Page**, who argue that fairness demands assigning burdens in proportion to contribution.

At first glance, the moral appeal of PPP is straightforward. Climate change is the cumulative result of greenhouse gas emissions that remain in the atmosphere for centuries. Since the industrial revolution, a small number of states—primarily in the Global North—have consumed a vastly disproportionate share of the atmospheric commons. The United States alone has emitted more than **400 GtCO$_2$**, roughly a quarter of all historical emissions; Europe has contributed a similar amount; China's cumulative emissions, by contrast, are less than half the U.S. total; India's are an order of magnitude smaller; many African countries have contributed almost nothing. If climate change is understood as a case of harm imposed on vulnerable populations, PPP seems to demand that those who damaged the climate system undertake the lion's share of mitigation and compensation.

Table 2 - Cumulative CO$_2$ Emissions 1850–2021 (approx.)

Country/Region	Cumulative Emissions (GtCO$_2$)	Share of Global Total	Per-Capita Historical Share*
United States	~410 Gt	~25%	>1,200 tonnes per person
European Union (27)	~350 Gt	~22%	~900–1,000 tonnes per person
China	~260 Gt	~14%	~180–200 tonnes per person
India	~60 Gt	~3%	<50 tonnes per person
Sub-Saharan Africa	~25 Gt	~1–2%	<20 tonnes per person

*Per-capita historical values are approximate, dividing cumulative emissions by today's population for rough comparison.

PPP also fits naturally with **Shue's distinction between "subsistence" and "luxury" emissions**. Whereas subsistence emissions are needed for basic human survival and cannot reasonably be condemned, luxury emissions—such as those associated with private aviation, meat-heavy diets, or excessive energy consumption—are neither necessary nor morally justified if they cause avoidable harm to others. Historically, most emissions in industrialised countries have been luxury emissions. PPP therefore assigns responsibility not only in proportion to quantity but also in proportion to moral relevance: the emissions most responsible for climate change are precisely those that were least necessary.

A further argument in favour of PPP concerns feasibility and transition. High historical emitters have also reaped the greatest **economic benefits** from fossil-fuel industrialisation. Modern wealth, infrastructure, and development trajectories in Europe and North America are intimately tied to centuries of carbon-intensive growth. PPP thus aligns with the intuition that it is unfair to require low-income countries—many of which remain energy-poor—to bear the same burdens as those who already benefited from the emissions that created the crisis.

Despite its normative appeal, PPP faces several significant challenges. The first concerns **involuntary or ignorant emissions**. Much historical pollution occurred before climate science identified the harmful effects of greenhouse gases. Caney and others have therefore argued for a "modified" PPP that assigns responsibility only for emissions undertaken with adequate knowledge or reasonable foreseeability of harm. Yet drawing this temporal boundary is difficult: knowledge of climate risks did not arrive suddenly but emerged gradually over decades. Moreover, even if some historical emitters were non-culpable in a moral sense, the *effects* of their emissions remain; the atmosphere cannot ignore CO_2 merely because it was emitted innocently. Many scholars therefore argue that PPP should be understood not as a principle of moral blame, but as a principle of **causal responsibility**: those who caused the harm should address it, regardless of intention.

A second critique concerns the **intergenerational dimension** of responsibility. Some defenders of PPP assign responsibility to present citizens for the actions of their ancestors on the grounds that they continue to benefit from inherited wealth, infrastructure, and institutions. Others object that it is unfair to ask current generations to bear full responsibility for actions they did not personally commit. These tensions raise deep questions about collective agency and intergenerational moral inheritance—questions that PPP alone cannot fully resolve.

A third challenge stems from **political feasibility**. As **John Broome** observes, PPP as well as the equal per-capita principle requires significant sacrifices from high-emitting countries—sacrifices that governments have repeatedly shown they are unwilling to make. The obligations implied by PPP would involve massive financial transfers from North to South and possibly legal responsibility for climate damages. Not surprisingly, the countries most responsible for historical emissions have consistently resisted binding commitments grounded in responsibility. Broome therefore argues that although PPP is morally compelling, it struggles to serve as the foundation of a politically workable climate regime.

These limitations do not undermine the moral force of PPP but reveal the need for complementary principles.

10.7 The Beneficiary-Pays Principle

The **beneficiary-pays principle (BPP)** offers an alternative route to assigning responsibility for climate change, complementing but also diverging sharply from the polluter-pays principle. Rather than grounding climate duties in *causal responsibility* for past emissions, BPP assigns obligations to those who **benefit from the processes that created the harm**, whether or not they were personally involved in causing it. The central moral claim is that **enjoying the advantages generated by unjust or harmful activities creates special duties to address their consequences**. In the climate context, this means that states and populations who continue to benefit from carbon-intensive development owe a proportionate share of the costs of mitigation, adaptation, and compensation.

The beneficiary-pays principle is closely associated with the work of **Henry Shue** and **Simon Caney**, who articulate complementary but distinct arguments for its adoption. Shue frames BPP as grounded in the enduring moral significance of **structurally embedded advantages**. In his influential essays on climate justice, Shue argues that today's high-income societies profit from centuries of fossil-fuel use: industrial infrastructure, transport systems, agricultural productivity, military security, and even public health capacity all emerged through carbon-intensive development. These benefits persist even if the original emitters acted without knowledge of the harms they were creating. What matters for Shue is that *the advantages remain*, and that these advantages were created through activities that now impose severe burdens—rising sea levels, droughts, floods—on others who did not share in their benefits. On this view, the persistence of unjust advantage suffices to generate compensatory duties.

Caney deepens and formalises this argument by distinguishing two independent bases for climate responsibility: **causal responsibility** and **beneficiary responsibility**. Even when agents are not causally responsible—because they are too young, because the original emissions occurred long ago, or because they lived in conditions of ignorance—they may still owe duties if they benefit from the harmful process. Caney's argument is particularly relevant for intergenerational justice. Present citizens of the United States or Europe may not have emitted the carbon that triggered global warming in the late nineteenth or early twentieth century, but they inherit societies whose economic and political structures rely on past emissions. As long as these benefits persist, Caney argues, so do the associated obligations. Beneficiary responsibility thus fills a conceptual gap left by the polluter-pays principle: it captures cases where the original wrongdoers are deceased or non-identifiable, yet the benefits and harms remain.

BPP also has implicit support in international law. The **UNFCCC's principle of common but differentiated responsibilities and respective capabilities (CBDR–RC)** acknowledges that the nations with the greatest capacity to act are overwhelmingly those that benefited most from historical emissions. Although CBDR stops short of naming BPP explicitly, the logic of inherited advantage permeates its structure, particularly in negotiations over finance, technology transfer, and adaptation.

Despite its normative appeal, the beneficiary-pays principle faces several important **challenges and critiques**. One difficulty concerns **the diffusion of benefits**. Industrial development, though uneven in distribution, has produced global spillovers: medicines, agricultural technologies, scientific knowledge, and communication systems that now benefit populations worldwide. Critics argue that because low-income countries also benefit from some products of high-carbon industrialization, BPP risks becoming indeterminate—if everyone benefits, assigning differential responsibilities becomes difficult. Defenders respond that while benefits are diffuse, *the overwhelming share of material, infrastructural, and financial gains has been captured by industrialized states*, thereby preserving differentiation.

A second critique focuses on **the fairness of inheriting moral responsibilities**. Not all beneficiaries of carbon-intensive development consent to their inherited advantages, nor did they necessarily endorse the emissions that produced them. Some theorists therefore argue that BPP risks unfairly burdening individuals who played no role in creating the problem. In reply, Shue and Caney contend that BPP is not about blame but about **fair burden-sharing**: even innocent beneficiaries may reasonably be asked to help remedy harms that stem from the systems that advantage them.

A third challenge concerns **political feasibility**. Like the polluter-pays principle, BPP frequently implies large financial transfers from wealthy to poorer countries. Yet political resistance in the Global North has repeatedly stalled proposals for climate finance, loss-and-damage funding, and technology transfer. Hence, while BPP is morally compelling, its institutional implementation remains limited—an issue that will become central when we consider Broome's argument that a just climate regime must ultimately align with self-interest through a **no-sacrifice** framework.

Finally, BPP struggles to account for states that are simultaneously **beneficiaries and victims**—notably China, Brazil, or South Africa—where industrial successes coexist with acute climate vulnerabilities. In these cases, beneficiary status does not map neatly onto capacity or culpability. This limitation highlights the need for complementary principles—such as the **ability-to-pay principle**, which allocates responsibilities based on economic capacity rather than inherited advantage, and Broome's **resource-constrained efficiency**, which reframes the problem around mutually beneficial transitions rather than moral sacrifice.

Thus, while the beneficiary-pays principle offers a powerful critique of inherited advantage and a compelling foundation for climate justice, it requires integration with broader distributive frameworks.

10.8 The ability-to-pay Principle

The ability-to-pay principle (ATP) offers a third major framework for allocating climate responsibilities. Unlike the polluter-pays principle, which grounds obligations in causal responsibility, or the beneficiary-pays principle, which appeals to inherited advantage, ATP assigns climate duties according to **economic capacity**. The central intuition is straightforward: those who are most able to bear the costs of mitigation, adaptation, and compensation should shoulder a correspondingly larger share of the burden. Climate change is a collective-action problem with uneven consequences; justice requires that the distribution of obligations be sensitive to disparities in wealth, technological capability, and fiscal resilience. *(see Caney 2010; Shue 1999)*

ATP has deep roots in distributive justice. Although **John Rawls** did not write directly about climate change, his principles of justice—particularly the **difference principle**—have been widely interpreted as supporting an ability-based allocation of climate burdens. *(Rawls 1971; Rawls 1999)* Rawls argues that inequalities are justified only if they benefit the least advantaged. Applied to the global climate regime, this suggests that affluent societies with extensive economic and institutional capacities should contribute more to global mitigation and adaptation, ensuring that climate governance improves the prospects of the world's poorest and most vulnerable communities. Several Rawlsian scholars, including **Henry Shue** and **Simon Caney**, build on this idea: if climate change threatens basic rights and undermines the conditions for a just global order, then those with the greatest capacity must take on the heaviest responsibilities. *(Shue 2014; Caney 2005)*

In policy terms, the ability-to-pay principle aligns with the logic of **progressive taxation**, a long-standing feature of public finance debates. *(Musgrave 1959)* Wealthier states have greater fiscal space to invest in renewable energy, to fund adaptation projects in vulnerable regions, and to absorb transitional costs associated with decarbonizing their economies. It is therefore efficient and fair, ATP proponents argue, to require them to do more. The principle also integrates smoothly with the "respective capabilities" clause of the UNFCCC's doctrine of **Common but Differentiated Responsibilities and Respective Capabilities (CBDR–RC)**. *(UNFCCC 1992; Rajamani 2012)* The Paris Agreement—even as it softened the interpretive force of historical responsibility—retained the claim that states' obligations should reflect their capacities.

The ability-to-pay principle is often presented as more **politically feasible** than strictly responsibility-based approaches. Whereas the polluter-pays principle (PPP) implies potentially vast reparative transfers, and the beneficiary-pays principle (BPP) carries deep historical and moral baggage, ATP is **forward-looking**: it does not require establishing blame or tracking inherited advantage. It simply requires allocating burdens in a way that reflects economic capability. For this reason, it is often the preferred principle among economists, international institutions, and policymakers seeking pragmatic solutions. *(Stern 2007; IPCC WGIII 2014)*

Yet ATP faces several significant **limits and critiques**.
A first critique is that ATP, when applied alone, may **disconnect responsibility from causation**. A wealthy country with low historical emissions—such as Norway—would owe much more under ATP than a lower-income but high-emitting country. Critics argue that this mismatch violates intuitions of **moral responsibility**: capacity alone should not determine liability for harm. *(Meyer & Roser 2010)* To address this, some theorists propose hybrid models combining ATP with PPP or BPP.

A second critique concerns **incentives and moral hazard**. If obligations are tied primarily to capacity, states may face less incentive to reduce emissions quickly if doing so affects their economic strength. Conversely, less affluent states might be implicitly rewarded for maintaining low levels of wealth, since poorer capacity reduces their obligations. While this objection is often overstated, it raises questions about ATP's long-term compatibility with global inequality. *(Gosseries 2004)*

A third critique highlights ATP's **insensitivity to historical injustice**. The principle is agnostic about how existing wealth was accumulated. Much of the wealth that gives affluent states their high capacity is itself a product of **centuries of carbon-intensive industrialization**. By ignoring causation and benefit, ATP risks allowing historically responsible states to fulfil their duties merely by virtue of being wealthy, without acknowledging the origins of their wealth. *(Roberts & Parks 2007)* In this sense, ATP may appear to excuse past exploitation.

A fourth critique concerns **domestic inequality**. ATP typically assesses states as collective agents, assigning obligations based on national income or GDP. Yet many high-capacity states contain significant populations who are economically vulnerable and who would bear disproportionate costs if national obligations were passed to them. Critics argue that ATP must therefore integrate some account of **within-country justice**, such as progressive domestic burden-sharing. *(Okereke 2010)*

Despite these objections, the ability-to-pay principle retains an important role in debates over global climate governance. It highlights the practical realities of fiscal and technological capacity, offering a clear and administratively tractable approach to distributing obligations. Moreover, ATP provides a compelling moral baseline: **those who can most easily prevent catastrophic harm ought to act first and most forcefully**. However, as the critiques suggest, ATP cannot stand alone as a comprehensive theory of climate justice. It must be situated within a broader framework that also recognizes historical responsibility, inherited advantage, and political feasibility.

In the next section, I turn to **Broome's no-sacrifice regime**, which attempts to combine the normative insights of these principles with institutional arrangements capable of overcoming the political impediments that have long prevented their realization. Broome's proposal reframes climate justice in terms of **resource-constrained efficiency** and the alignment of climate action with self-interest—a crucial step in bridging the gap between moral theory and workable global policy. *(Broome 2019; Broome 2020)*

10.9 Broome's No-Sacrifice Principle – From Externality to Institutional Design

John Broome's no-sacrifice framework begins from a double diagnosis: **the international effort to control climate change is failing**, and it is failing **because we**

have tried to motivate it primarily through morality. Governments applauded the Paris Agreement's goal of keeping warming well below 2°C and "pursuing efforts" towards 1.5°C, but within days key states were rolling back support for renewables, opening new fossil fuel reserves, and relaxing incentives for low-emission technologies. Pledges remain collectively insufficient; many countries are not even on track to meet their own targets. The Australian case—high per-capita emissions, coal exports, creative accounting with "Kyoto credits", and the dismantling of a carbon tax while the country burns—serves as a particularly stark example of how governments are willing to ignore both science and public opinion when powerful fossil interests are at stake.

Broome insists that the **moral critique itself is correct**. Climate change is an injustice, inflicted by some people—especially the old, rich and high-emitting—on others, particularly the young and the poor. The moral voice of Greta Thunberg and the language of intergenerational responsibility are entirely justified. The problem is not that morality is mistaken, but that **morality alone is too weak** to move governments that are structurally captured by short-term interests. Some individuals restrict their own emissions out of conscience, but large-scale decarbonisation requires structural change that only states can implement, and many states are, in practice, largely impervious to ethical exhortation when it conflicts with powerful lobbies. Broome's central strategic conclusion is therefore blunt: if we continue to design climate policy on the assumption that present generations will voluntarily **sacrifice** for the future, we will fail.

From this conclusion he pivots to a different starting point: **self-interest**, backed by economic theory. The key conceptual move is to treat greenhouse gas emissions as what economists call an **externality**. When an agent burns fossil fuel, the resulting greenhouse gases disperse globally and cause small harms, spread over billions of people and many decades. These harms are part of the real cost of the activity, but the emitter does not bear them; they fall on others. This is the classical structure of an externality.

In standard welfare economics, this leads to a familiar logical chain:

1. **Emissions are an externality.**
2. **Externalities cause Pareto inefficiency**: the market outcome is not Pareto-optimal.
3. In a **Pareto-inefficient** state, there exists, in principle, a policy change that makes at least one person better off **without making anyone worse off**.

So, in such a framework, there should exist a way of correcting the emissions externality such that **no one needs to sacrifice**. Broome initially endorsed this textbook conclusion: if climate change is "just" an externality, then a suitably designed policy could in principle make everyone at least as well off as before, some better off, and avoid any losers.

However, he then shows why this is **not quite true** for climate change. The textbook Pareto result assumes a **fixed population**: the same people exist under all policy options. But climate policy is intergenerational in a strong sense. Serious mitigation will change how and where people live, how often they travel, how cities are organised, and so on. Over time, this alters who meets whom, who has children with whom, when they have children, and how many children they have. Different climate policies produce **different future people**. In such a setting we cannot say that there exists some policy that makes "all" individuals, including all future individuals, at least as well off as they would otherwise have been, because the set of individuals differs between policies. The standard Pareto theorem simply does not apply.

Rather than abandon the idea of no-sacrifice, Broome **repairs the efficiency concept**. He shifts from welfare comparisons across all present and future individuals to a resource-based criterion. He proposes **resource-constrained efficiency**: a situation is inefficient if there is a feasible change that

- makes all **currently existing** people at least as well off as before (and some better off), and
- leaves the **stock of resources** available to future people at least as good as it would otherwise have been.

This avoids the non-identity problem by not requiring that any particular future individual be better off; it only requires that future generations, whoever they are, inherit resources—capital, infrastructure, environmental quality—no worse than under the status quo. Broome argues that intergenerational externalities like greenhouse gases generally render the world **resource-inefficient** in this sense. Thus, in principle, there is a way to control climate change that **does not make any currently living group worse off**, while leaving future generations with at least as good resources. In that precise sense, **climate mitigation can be no-sacrifice**.

The challenge is then institutional: how do we design such a policy? Broome's answer is a package with four core components:

1. **Carbon pricing (typically a carbon tax).**
2. **Compensation** to those who would otherwise lose from the tax and from decarbonisation.
3. **Government borrowing** to shift part of the costs to future beneficiaries.
4. A new **World Climate Bank** plus **enforcement mechanisms** against free riders.

First, an externality must be **internalised**. A carbon tax equal to the marginal external cost of emissions makes emitters face the true social cost of their activities. This drives the economy towards efficiency by discouraging carbon-intensive production and consumption and encouraging cleaner alternatives. However, such a tax clearly **hurts some present agents**: households with carbon-heavy consumption, workers in coal, oil and gas sectors, and firms whose assets are tied to fossil fuels. Left on its own, carbon pricing creates losers and so provokes political resistance.

To preserve the no-sacrifice condition, Broome proposes that governments **recycle carbon tax revenue** and supplement it with **borrowing**. Carbon tax receipts can be returned through reductions in other taxes (such as income tax), lump-sum transfers, or targeted support for vulnerable groups. In addition, governments can issue long-term bonds whose repayment obligations fall on future taxpayers. These borrowed resources can be used to finance generous **compensation schemes**—for example, subsidies and transition packages for fossil-fuel workers and households, or buy-outs for shareholders in carbon-intensive industries. The aim is to arrange matters so that each major class of present agents—workers, consumers, firms, states—is at least as well off as they would have been without climate policy, even while emissions are driven down.

A similar logic applies to **investment**. Currently, much of global investment flows into conventional capital—roads, buildings, fossil-fuel infrastructure—while we pass on to future generations a dangerously degraded atmosphere. Broome's idea is not to increase total sacrifice today, but to **change the composition** of what we pass forward: less dirty capital and more green capital. By borrowing and redirecting investment into renewables, energy efficiency, and clean transport, we leave future people a different bundle of resources: a cleaner atmosphere plus a stock of low-carbon infrastructure. Because future generations benefit from avoided damages and improved capital, they can justifiably bear a portion of the costs through servicing the debts we incur today.

This is how the **intergenerational transfer** works in his framework. Climate mitigation, if successful, generates very large net benefits in the future. A no-sacrifice regime, instead of allowing all those benefits to accrue to future people, "pulls some of them back" to the present through borrowing: we take some of the future surplus in the form of present consumption and compensation, while still leaving future generations better off than they would have been under unmitigated climate change.

Yet this financing strategy faces real-world constraints. Many governments are already heavily indebted or constrained by austerity doctrines, and poorer states cannot borrow cheaply. To overcome this, Broome proposes a new international financial institution: a **World Climate Bank**. Modelled on the World Bank created after the Second World War, it would be backed by the credit of fiscally strong states, issue very long-term bonds, and lend to countries needing to finance green investment and compensation. Given that the benefits of mitigation unfold over many decades, extremely long maturities—even perpetual bonds—fit the temporal structure of the problem. A World Climate Bank would make a no-sacrifice transition **financially accessible** even to countries currently excluded from affordable capital markets.

One major obstacle remains: **free riding**. Even if a no-sacrifice policy is collectively beneficial, each individual state has an incentive to let others undertake the effort while continuing business as usual. Broome likens this to a prisoner's dilemma. Some governments, such as Australia, explicitly adopt this stance: since their unilateral emission cuts have negligible impact on their local climate, they prefer to rely on others' mitigation while continuing to export and burn fossil fuels. For a no-sacrifice system to work, participating countries must develop **defences against free riders**—for example, punitive tariffs or border carbon adjustments on imports from non-cooperating states. In addition, Broome concedes that **moral pressure still has a**

role here: even if we abandon the hope that pure morality can drive large-scale sacrifice, social movements and public opinion can still stigmatise free-riding behaviour and push governments toward cooperation.

Broome is explicit about the **normative cost** of his approach. A no-sacrifice regime does not correct either global inequality or climate injustice. Those who profited from past emissions are, in effect, paid to stop harming others, and their advantaged position is preserved. As he notes, this looks uncomfortably like paying a bully to stop bullying: the harm ceases, but the injustice is not rectified. Moreover, the no-sacrifice constraint rules out using climate policy as a vehicle for major redistributive reforms; inequality must be addressed by other means. Nonetheless, Broome argues, these costs are **forced upon us** by the long-running failure of the moral-sacrifice strategy. A world that continues on a high-emissions path in the name of ideal justice is, in practice, worse than a world that limits warming through a second-best but feasible no-sacrifice regime.

In this sense, Broome's no-sacrifice framework is best understood as a **political-economic supplement** to the more overtly moral principles discussed earlier—polluter pays, beneficiary pays, and ability to pay. It does not deny their moral force, but it claims that those principles, by themselves, cannot overcome entrenched interests and short-termism. Only by **aligning climate action with the self-interest of all powerful actors**—through carbon pricing, compensation, borrowing, international climate finance, and credible sanctions against free riders—do we have a realistic hope of controlling climate change.

This framework has several **appealing features**. First, it collapses the political resistance that often accompanies responsibility-based principles: fossil-fuel interests, energy-intensive industries, and wealthy states would not face uncompensated losses. Second, it reframes climate mitigation as a win-win investment problem rather than a burden-sharing conflict. Third, it acknowledges that climate change is fundamentally an intergenerational problem: the key challenge is aligning the incentives of present agents with the interests of future ones.

However, Broome's approach also faces important **limits and critiques**.

The Justice Objection

Broome acknowledges that the no-sacrifice regime **fails to rectify injustice**. Under this scheme, historical emitters are effectively *paid* to stop causing harm, in the same way that a bully might be paid to stop bullying. This sits uneasily with principles of corrective justice. The approach leaves intact the unjust advantages accumulated through centuries of carbon-intensive development and requires victims—future generations—to finance compensation to present beneficiaries. Critics argue that this amounts to institutionalising climate injustice rather than overcoming it. *(Shue 2014; Meyer & Roser 2010)*

The Maldistribution Objection

Many theorists worry that a no-sacrifice approach **perpetuates global economic inequality**. Since no one may be made worse off, the regime cannot redistribute resources from rich to poor—even when doing so would produce a more just distribution. Broome concedes that the optimal distribution of income or wealth may be unattainable under the no-sacrifice constraint. Climate policy alone would not correct global inequality, even though inequality was produced in part by fossil-fuel development. Critics argue that climate justice cannot be divorced from economic justice, and that Broome's model fragments the two problems artificially. *(Caney 2010; Roberts & Parks 2007)*

The Feasibility Objection

The practical success of Broome's proposal depends on **large-scale public borrowing**, including in lower-income countries with limited creditworthiness. Many states face debt constraints, austerity norms, or pressures from international financial institutions that make additional borrowing politically or legally difficult. Broome proposes the creation of a **World Climate Bank**, designed to lend on the basis of the credit of wealthy states, but establishing such an institution would require unprecedented geopolitical cooperation. Critics argue that the political barriers Broome seeks to escape may simply reappear in another form. *(Pogge 2016; Bauer & Richards 2021)*

The Free-Riding Problem

Even under a no-sacrifice regime, climate change remains a global prisoner's dilemma. A state may find it rational to enjoy the benefits of others' mitigation efforts while doing little itself. Broome suggests tariffs and global public pressure as tools for discouraging free-riding, but he does not fully resolve the structural incentive problem inherent in global climate governance. *(Barrett 2003)*

The Normative Thinness Objection

Some critics argue that Broome's no-sacrifice approach—while politically pragmatic—abandons the deeper moral content of climate justice. By redefining justice in terms of resource-constrained efficiency rather than responsibility, rights, or equity, the framework risks reducing climate ethics to a purely economic optimisation problem. Principles such as PPP, BPP, and the ability-to-pay principle capture morally salient relationships of harm, benefit, and capacity that the no-sacrifice framework sidelines. *(Gardiner 2011)*

Despite these objections, Broome's contribution is invaluable. It highlights that **moral principles alone cannot generate political action**, and that climate institutions must be designed with attention to self-interest, incentives, and economic feasibility. The no-sacrifice framework does not replace responsibility-based principles, but reframes them: even morally compelling duties cannot motivate action unless they are embedded within institutions that align justice with self-interest. Broome's proposal challenges theorists to think creatively about institutional design and to consider how climate justice can be achieved *in the world as it is*, rather than only as it ought to be.

10.10 Conclusion: Choosing Under Pressure

This chapter has argued that climate change confronts us with two tightly entangled but conceptually distinct tasks. The first is **economic**: to understand the nature of the climate problem as a market failure and to identify efficient instruments—such as carbon pricing and cap-and-trade—that can correct it at lowest cost. The second is **moral and political**: to decide how the burdens and benefits of climate action should be shared among countries and generations in a way that is recognisable as just.

On the economic side, the picture is clearer than public debate sometimes suggests. Once we treat greenhouse gas emissions as a pervasive externality, the logic of welfare economics points straight toward putting a price on carbon. Whether via a tax or an emissions trading system, the goal is to align **marginal private cost** with **marginal social cost**, moving global emissions from an inefficiently high level toward a socially efficient one. The Stern Review, and subsequent work in climate economics, demonstrates that when we take seriously two central features of the problem—**deep uncertainty** and the **long-lived nature of CO_2**—the case for strong, early mitigation is robust across a wide range of assumptions.

Even for those who place a relatively low weight on the future, the arithmetic of small changes in growth rates compounding over centuries, combined with the non-negligible risk of catastrophic outcomes, makes it economically rational to pay a modest cost now to avoid very large damages later. Climate policy, on this view, is less a heroic sacrifice for distant generations than a form of long-term insurance: a premium we pay to reduce the probability of outcomes that would be intolerable, both materially and morally.

At the same time, we saw that this economic clarity does not settle the **distributional question**. Efficiency tells us that global emissions must fall, but it does not tell us who should cut how much, who should pay for adaptation, or how responsibility for past and future emissions should be divided. Here we step into the domain of climate justice.

We examined five leading principles that structure contemporary debates:

- **Equal per-capita emissions**, which treats the atmosphere as a shared global commons and argues that all persons have an equal claim to its use.
- The **polluter-pays principle**, which assigns burdens in proportion to causal responsibility for emissions.
- The **beneficiary-pays principle**, which focuses instead on those who continue to benefit from carbon-intensive development, whether or not they personally caused the harm.
- The **ability-to-pay principle**, which distributes obligations according to economic capacity, emphasising feasibility and the protection of the least advantaged.
- **Broome's no-sacrifice framework**, which seeks to redesign institutions—through carbon pricing, compensation, borrowing, and a World Climate Bank—

so that mitigation can, in principle, proceed without making any present group worse off.

Each of these principles captures something morally important. Equal per-capita rights express the basic idea of moral equality; polluter pays reflects intuitions about responsibility and corrective justice; beneficiary pays highlights the significance of inherited advantage; ability to pay speaks to solidarity and the moral relevance of capacity; and the no-sacrifice approach forces us to confront the limits of exhortation and the need to align justice with self-interest.

Yet each principle also faces serious objections. Equal per-capita allocations, if applied strictly, would require changes so radical from high-emitting countries that they appear politically unattainable. Polluter pays struggles with historical ignorance, intergenerational responsibility, and the fact that many of the original emitters are long dead. Beneficiary pays raises difficult questions about diffuse benefits and the fairness of imposing obligations on innocent heirs. Ability to pay can disconnect responsibility from causation and underplay historical injustice. Broome's no-sacrifice regime, while politically astute, leaves much injustice untouched and depends on ambitious institutional innovations that are far from guaranteed.

These tensions mirror the broader dilemma that runs through the chapter: the uneasy relationship between **justice and feasibility**, or, more starkly, between justice and the risk of catastrophic failure. If we insist too rigidly on ideal fairness—on a single pure principle, applied without compromise—we may never reach an agreement capable of stabilising the climate. If, by contrast, we abandon justice entirely in favour of what is easiest politically, we risk entrenching historical inequalities and legitimising a profoundly unequal distribution of climate harms. In either case, the greatest losses would fall on those with the least power to shape the outcome: future generations and today's poorest and most vulnerable.

This is where the current shift in many governments' attitudes toward climate policy becomes especially troubling. As geopolitical tensions rise, energy security concerns sharpen, and populist narratives gain traction, it becomes tempting for states to retreat into a narrow, short-term calculus of national interest. Climate targets are delayed, diluted, or reframed; investments in fossil infrastructure are extended; and the language of ambition coexists uneasily with decisions that lock in higher emissions for decades. These reversals do not only undermine efficiency; they corrode trust, making it harder to sustain any conception of fair burden-sharing.

The central lesson of this chapter is not that we must choose between economics and ethics, or between ideal justice and crude self-interest. Rather, it is that **durable climate policy must be built at their intersection**. Economic analysis helps us see that ambitious mitigation is not a utopian demand but a rational response to a severe, long-lived risk. Normative principles help us see that who pays and who benefits matters, not only for philosophical coherence but for political legitimacy. Institutional design—whether in the form of carbon pricing, cap-and-trade systems, climate finance, or a possible World Climate Bank—must be guided by both sets of insights.

The “climate choice” is therefore not a single decision, taken once and for all at Paris or Glasgow, but an ongoing process of negotiation over **how much risk we are willing to pass on** and **how fairly we are prepared to share the effort of reducing it**. In a world where some governments are stepping back from their earlier commitments, the temptation is either to lower our sights or to retreat into moral indictment alone. This chapter has suggested a different response: to take seriously both the economic structure of the problem and the plurality of plausible principles of justice, and to use them together to design policies that are fair enough to command support and effective enough to matter.

Whether we succeed will not be determined by theory alone. It will depend on political movements, technological change, institutional innovation, and the willingness of societies to see beyond the electoral cycle. But a clearer understanding of the economic and ethical contours of the climate problem can at least help us recognise what is at stake. The choice is not between action at a fair price and inaction at no cost. It is between imperfect but workable paths toward a livable future and a trajectory that leaves future generations with a hotter, more unequal, and more dangerous world than justice—or prudence—can accept.

Bibliographical References:

Appiah, Kwame Anthony. (2005). *The ethics of identity*. Princeton University Press.

Aristotle. *The Nicomachean Ethics*. Translated by J. A. K. Thomson, revised by Hugh Tredennick and Jonathan Barnes. London: Penguin Books, 2004.

Arrow, Kenneth J. (1951). *Social choice and individual values*. Yale University Press.

Arrow, Kenneth J. (2007). Global climate change: A challenge to policy. *Economists' Voice*, 4(3).

Arrow, Kenneth J., Cropper, Maureen, Gollier, Christian, Groom, Ben, Heal, Geoffrey, Newell, Richard, Nordhaus, William, Pindyck, Robert, Sterner, Thomas, Tol, Richard, & Weitzman, Martin. (2013). Determining benefits and costs for future generations. *Science*, 341(6144), 349–350.

Barrett, Scott. (2003). *Environment and statecraft: The strategy of environmental treaty-making*. Oxford University Press.

Bauer, Michael, & Richards, Kenneth. (2021). Climate finance and debt constraints. *Climate Policy*, 21(7), 874–889.

Bentham, Jeremy. (1996). *An introduction to the principles of morals and legislation* (J. H. Burns & H. L. A. Hart, Eds.). Oxford University Press. (Original work published 1789)

Borjas, George J. (2015). Immigration and globalization: A review essay. *Journal of Economic Literature*, 53(4), 961–974.

Borjas, George J., & Monras, Joan. (2016). The labour market consequences of refugee supply shocks. *Economic Policy*, 31(87), 497–550.

Bostrom, Nick**.** (2002). *Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards*. Journal of Evolution and Technology 9: 1–30.

Broome, John. (1992). *Counting the cost of global warming*. White Horse Press.

Broome, John. (2012). *Climate matters: Ethics in a warming world*. W. W. Norton.

Broome, John. (2019). Resource-constrained efficiency and climate economics. *Oxford Review of Economic Policy*, 35(3), 439–457.

Broome, John. (2020). Self-interest against climate change. *Environmental Values*, 29(1), 1–20.

Caney, Simon. (2005). *Justice beyond borders*. Oxford University Press.

Caney, Simon. (2010). Climate change and the duties of the advantaged. *Critical Review of International Social and Political Philosophy*, 13(1), 203–228.

Carens, Joseph H. (1987). Aliens and citizens: The case for open borders. *The Review of Politics*, 49(2), 251–273.

Carens, Joseph H. (2013). *The ethics of immigration*. Oxford University Press.

Clemens, Michael A. (2011). Economics and emigration: Trillion-dollar bills on the sidewalk? *Journal of Economic Perspectives*, 25(3), 83–106.

Clemens, Michael A. (2014). Does development reduce migration? In *International Handbook on Migration and Development*. Edward Elgar.

Clemens, Michael A., Montenegro, Claudio, & Pritchett, Lant. (2019). The place

premium. *Review of Economics and Statistics*, 101(2), 201–213.
Cranston, Maurice. (1967). *John Stuart Mill: The man*. Allen & Unwin.
Epicurus. *Letter to Menoeceus*. In Diogenes Laertius, *Lives of Eminent Philosophers*. (Standard Loeb edition.)
Friedman, Milton. (1962). *Capitalism and freedom*. University of Chicago Press.
Fukuyama, Francis. (2018). *Identity: The demand for dignity and the politics of resentment*. Farrar, Straus and Giroux.
Gandhi, Mohandas Karamchand. (1968). *Ethical Religion*. Ahmedabad: Navajivan Trust.
Gardiner, Stephen. (2011). *A perfect moral storm: The ethical tragedy of climate change*. Oxford University Press.
Gosseries, Axel. (2004). Historical emissions and free riding. *Ethical Perspectives*, 11(1), 36–60.
Goulder, Lawrence, & Schein, Andrew. (2013). Carbon taxes vs. cap-and-trade. *Climate Change Economics*, 4(3).
Hart, Herbert Lionel Adolphus. (1970). *Bentham*. Methuen.
Hayek, Friedrich August von. (1960). *The constitution of liberty*. Routledge.
Heal, Geoffrey, & Millner, Alex. (2014). Uncertainty and decision making in climate change economics. *Review of Environmental Economics and Policy*, 8(1), 120–137.
Hicks, John R. (1939). *Value and capital*. Oxford University Press.
IPCC. (2014). *Mitigation of climate change* (WGIII contribution to AR5).
IPCC. (2021). *Sixth Assessment Report: The physical science basis*.
Jaggar, Alison M., Edited, (2010). *Thomas Pogge and His Critics*. Polity.
Kant, Immanuel. (2012). *Groundwork of the metaphysics of morals* (M. Gregor & J. Timmermann, Trans.). Cambridge University Press. (Original work published 1785)
Kant, Immanuel. "Perpetual Peace: A Philosophical Sketch." In *Kant: Political Writings*, edited by Hans Reiss, translated by H. B. Nisbet, 93–130. Cambridge: Cambridge University Press, 1991.
Kass, Leon. (2002). *Life, liberty, and the defense of dignity*. Encounter Books.
Marshall, Alfred. (1890). *Principles of economics*. Macmillan.
Marshall, Thomas Humphrey. (1950). *Citizenship and social class*. Cambridge University Press.
Meyer, Lukas, & Roser, Dominic. (2010). Climate justice and historical emissions. *Critical Review of International Social and Political Philosophy*, 13(1), 229–253.
Mill, John Stuart. (1859/1978). *On liberty* (D. L. Coleman, Ed.). Penguin.
Mill, John Stuart. (1861/2001). *Utilitarianism* (R. Crisp, Ed.). Oxford University Press.
Milanović, Branko. (2016). *Global inequality*. Harvard University Press.
Musgrave, Richard. (1959). *The theory of public finance*. McGraw-Hill.
Nagel, Thomas. (1979). Death. In *Mortal Questions*. Cambridge University Press.

Nagel, Thomas. (2005) *The Problem of Global Justice.* Philosophy & Public Affairs 33, no. 2: 113-147.
Nordhaus, William. (2008). *A question of balance*. Yale University Press.
Nozick, Robert. (1974). *Anarchy, state, and utopia*. Basic Books.
Okereke, Chukwumerije. (2010). Climate justice and the politics of scale. *Cambridge Journal of Regions, Economy and Society*, 3(2), 189–205.
Pareto, Vilfredo. (1906). *Manual of political economy*. (Standard English translation 1971, A.M. Kelley.)
Parfit, Derek. (1984). *Reasons and persons*. Oxford University Press.
Pigou, Arthur Cecil. (1920). *The economics of welfare*. Macmillan.
Pindyck, Robert. (2013). Climate change policy: What do models tell us? *Journal of Economic Literature*, 51(3), 860–872.
Pogge, Thomas. (2002). *World poverty and human rights*. Polity Press.
Pogge, Thomas. (2016). Global justice and carbon responsibilities. In *The Oxford Handbook of International Climate Change Law*.
Pritchett, Lant. (2006). *Let their people come*. Center for Global Development.
Pritchett, Lant. (2022). *Workers without borders* (Working paper). Kennedy School.
Pritchett, Lant, & Clemens, Michael. (2019). Labor mobility as a win–win policy. *Economic Record*, 95(310), 1–17.
Rajamani, Lavanya. (2012). The changing fortunes of CBDR in the UNFCCC. *International and Comparative Law Quarterly*, 61(3), 501–524.
Rawls, John. (1971). *A theory of justice*. Harvard University Press.
Rawls, John. (1999). *The law of peoples*. Harvard University Press.
Roberts, J. Timmons, & Parks, Bradley. (2007). *A climate of injustice*. MIT Press.
Rodrik, Dani. (1994). Political economy of trade policy. In G. Grossman & K. Rogoff (Eds.), *Handbook of international economics* (Vol. III). North-Holland.
Rodrik, Dani. (1997). *Has globalization gone too far?* Institute for International Economics.
Rodrik, Dani. (2017). Is global inequality the enemy of national equality? NBER Working Paper 23366.
Rothbard, Murray Newton. (1973). *For a new liberty*. Macmillan.
Rothbard, Murray Newton. (1982). *The ethics of liberty*. Humanities Press.
Ruhs, Martin. (2013). *The price of rights*. Princeton University Press.
Ruhs, Martin. (2019). On the ethics of labour immigration policy. In S. Fine & L. Ypi (Eds.), *Migration in political theory*. Oxford University Press.
Sandel, Michael J. (2007). *The case against perfection*. Harvard University Press.
Schneewind, Jerome B. (1990). *The invention of autonomy*. Cambridge University Press.
Sen, Amartya. (1970). The impossibility of a Paretian liberal. *Journal of Political Economy*, 78(1), 152–157.
Sen, Amartya. (1999). *Development as freedom*. Oxford University Press.
Sen, Amartya. (2009). *The Idea of Justice*. Belknap Press.

Shue, Henry. (1980). *Basic rights*. Princeton University Press.
Shue, Henry. (1992). Global environment and international inequality. *International Affairs*, 68(3).
Shue, Henry. (1993). Subsistence emissions and luxury emissions. *Law & Policy*, 15(1).
Shue, Henry. (2014). *Climate justice*. Oxford University Press.
Singer, Peter. (1972). Famine, affluence, and morality. *Philosophy & Public Affairs*, 1(3), 229–243.
Singer, Peter. (1975). *Animal liberation*. HarperCollins.
Singer, Peter. (2002). *One world*. Yale University Press.
Singer, Peter. (2009). *The life you can save*. Random House.
Singer, Peter. (2011). *Practical ethics* (3rd ed.). Cambridge University Press.
Singer, Peter. (2015). *The most good you can do*. Yale University Press.
Smith, Adam. (1759/1976). *The theory of moral sentiments*. (Standard Glasgow Edition.)
Smith, Adam. (1776/1976). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Edited by R. H. Campbell and A. S. Skinner. Glasgow Edition. Oxford.
Stern, Nicholas. (2007). *The economics of climate change: The Stern review*. Cambridge University Press.
Stiglitz, Joseph, & Stern, Nicholas. (2017). *Report of the High-Level Commission on Carbon Prices*. World Bank.
Taylor, Charles. (1992). The politics of recognition. In A. Gutmann (Ed.), *Multiculturalism: Examining the politics of recognition*. Princeton University Press.
UNFCCC. (1992). *United Nations Framework Convention on Climate Change*.
Weil, Simone. (1949/2002). *The need for roots*. Routledge.
Weitzman, Martin L. (2009). On modeling and interpreting the economics of catastrophic climate change. *Review of Economics and Statistics*, 91(1), 1–19.
Wellman, Christopher Heath. (2008). Freedom of association and the right to exclude. *Ethics*, 119(1), 109–141.
Walzer, Michael. (1983). *Spheres of justice*. Basic Books.
Walzer, Michael. (1994). *Thick and Thin: Moral Argument at Home and Abroad*. University of Notre Dame Press.